

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:06:48 ; Search time 22.0758 Seconds
(without alignments)
1638.789 Million cell updates/sec

Title: US-10-672-399-4
Perfect score: 1987
Sequence: 1 MRQDAPKPTPAARCSGLA.....TRRWSPACSPSSWRPLCCVC 376

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1426	71.8	543	2 JC4812	hyaluronan synthas
2	802.5	40.4	588	2 A43740	DG42 protein - Afr
3	308	15.5	568	2 T17588	hyaluronoglucosami
4	213	10.7	424	1 ZZRCL	nodulation protein
5	208.5	10.5	426	2 A95321	NODC N-ACETYLGLUCO
6	205.5	10.3	426	1 ZZRCL	nodulation protein
7	199.5	10.0	413	1 S34305	nodulation protein
8	176.5	8.9	424	1 S12793	nodulation protein
9	163.5	8.2	1086	2 JG6079	chitin synthase (E
10	159.5	8.0	395	1 JQ0396	nodulation protein
11	159	8.0	1195	2 S61886	chitin synthase (E
12	156.5	7.9	1165	1 S45879	chitin synthase (E
13	153.5	7.7	1239	2 T42020	class IV chitin sy
14	151.5	7.6	1112	2 T30202	probable chitin sy
15	148	7.4	419	2 A33100	hyaluronate syntha
16	146	7.3	395	2 A48755	hyaluronan synthas
17	145.5	7.3	1041	2 T31097	chitin synthase (E
18	142	7.1	1175	2 S39951	chitin synthase (E
19	140	7.0	743	2 T34632	probable bi-functi
20	137.5	6.9	428	1 E38180	nodulation protein
21	133.5	6.7	366	2 E59102	hypothetical prote
22	123.5	6.2	1103	2 T42022	probable chitin sy
23	119.5	6.0	1198	2 T28678	glycosyltransferas
24	115.5	5.8	447	2 A97211	chitin synthase (E
25	114.5	5.8	1498	2 S78102	chitin synthase (E
26	114.5	5.8	1852	2 JCS546	chitin synthase (E
27	114.5	5.8	1869	2 A59290	class V chitin syn
28	111.5	5.6	412	2 B90075	intercellular adhe
29	111	5.6	534	2 T35190	hypothetical prote

30 111 5.6 676 1 WMBEX6 UL6 protein - huma
31 108.5 5.5 869 2 H83500 probable glucosyl
32 106 5.3 487 2 S61243 deoxyribonuclease
33 106 5.3 3519 2 S43048 polyketide synthas
34 105 5.3 615 2 E70663 probable ppg prote
35 103 5.2 3573 2 E23070 erythronolide synt
36 100.5 5.1 3034 2 T14119 seven-pass transme
37 100 5.0 1155 2 AC2675 chromosome segrega
38 100 5.0 1165 2 A97457 structural mainten
39 99.5 5.0 291 2 B82645 phenylacetaldehyde
40 99.5 5.0 412 2 S77608 probable intercell
41 99.5 5.0 515 2 AC2635 glycosyltransferas
42 99.5 5.0 645 2 B97417 probable membrane
43 99 5.0 352 2 F75099 rhamnosyl transfer
44 99 5.0 6420 2 T30283 polyketide synthas
45 98.5 5.0 2314 2 T28698 hypothetical prote

ALIGNMENTS

RESULT 1

JC4812
hyaluronan synthase (EC 2.4.1.1) - human
C:Species: Homo sapiens (man)
C>Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 02-Mar-2001
C:Accession: JC4812
R:Itano, N.; Kimata, K.
Biochem. Biophys. Res. Commun. 222, 816-820, 1996
A:Title: Molecular cloning of human hyaluronan synthase.
A:Reference number: JC4812; MUID:96244584; PMID:8651928
A:Accession: JC4812
A:Molecule type: mRNA
A:Residues: 1-543 <ITA>
A:Cross-references: DDBJ:D84424; NID:gl401033; PIDN:BAAL2351.1; PID:d1013030; PID:gl401030
C:Comment: This enzyme synthesizes a high molecular weight glycosaminoglycan composed of
C:Genetics:
A:Gene: GDB:HAS1; HAS
A:Cross-references: GDB:L220109; OMIM:601463
A:Map position: 19q13.4-19q13.4
C:Keywords: Glycosyltransferase; hexosyltransferase; phosphoprotein
F:58,167,308,321/Binding site: phosphate (Thr) (covalent) #status predicted
F:82,247/Binding site: phosphate (Ser) (covalent) #status predicted

HA1a

RESULT 2

A43740
DG42 protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)

Query Match	10.3%;	Score	205.5;	DB	1;	Length	426;
Best Local Similarity	25.7%;	Pred. No.	1.2e-08;				
Matches	80;	Conservative	31;	Mismatches	101;	Indels	99;
Gaps	10;						
QY	49	DRYGLAFLGYGAFLSAHLVAQSLFAYLEHRRV-AAAARGPLDAATARSVALTISAYQE	106				
DB	5	DTTSTAISITAYULLTAYRSMQVLYA-----RPIDGPAVSAEPYETREPLPAVDIVPFSFNE	60				
QY	107	DPAYLQCLASARALLYPRARLRVLMVDGNRAEDLYMDFREVFADDPATYVMDGNY	166				
DB	61	DPGILSACLASIAQDQYP-GELRVYVDDGSRNR-----	93				
QY	167	HQPWEPAAGAVGAGAYREVEAEDPGCLAVEALVTRRCVCVCAQRMG-----GKREV	218				
DB	94	-----EAVRVRAFYSRDPFRFSFILLPENVGKKA	123				
QY	219	MYTAFKALGDSV-DYVQVCDSDTRLDPMALLELVRVLDEDPRGVAGGVDVRIINPLDSDW	277				
DB	124	Q-----IAAIGQSSGDLVNLVDSDTIAFDVVVSKLAKM-RDPEYGVAVGQLTASNSGDTWL	179				
QY	278	SFLSSRLYVAFNVNVERACQSYFHCVSCISGSLGTPPGPAATQRRP-----	322				
DB	180	TKLIDMEYLACNEERRAAQSRFGAVMCCC-----GPCMYRRSALASLLDQYETQLFR	232				
QY	323	-RPSCGGASRH	332				
DB	233	GKPSDFGEDRH	243				
RESULT 7							
S34305							
modulation protein nodC - Rhizobium sp.							
C:Species: Rhizobium sp.							
C:Date: 13-Jan-1995	#sequence_revision	27-Jan-1995	#text_change	09-Jul-2004			
C:Accession: S34305							
R;Relix, B.; Perret, X.; Golinowsky, W.; Pueppke, S.G.; Krishnan, H.B.; Brought							
submitted to the EMBL Data Library, June 1993							
A:Description: Lipo-oligosaccharide Nod-factor signals permit rhizobial penetra							
A:Reference number: S34303							
A:Accession: S34305							
A:Molecule type: DNA							
A:Residues: 1-413	<REL>						
A:Cross-references: UNIPROT:P50357; EMBL:X73362; NID:g312347; PIDN:CAAS1774.1;							
C:Comment: This is one of the proteins, coded by modulation genes, that are req							
C:Genetics:							
A:Gene: nodC							
C:Superfamily: modulation protein nodC							
C:Keywords: modulation							
Query Match	10.0%;	Score	199.5;	DB	1;	Length	413;
Best Local Similarity	26.5%;	Pred. No.	3.3e-08;				
Matches	81;	Conservative	28;	Mismatches	102;	Indels	95;
Gaps	11;						
QY	52	GLLAFLGYGAFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATA-RSVALTISAYQEDPAY	110				
DB	8	GAVAISIXAALSTAYKGMQAIYALPTN---TTAASTPVTGSGAPPSVDVIVPCVNEPDRA	64				
QY	111	LROCLASARALLYPRARLRVLMVD-----AEDLYMVD-MFREVFADDPATYVW	162				
DB	65	LSACLASIAQDY-AGELRVYVVDDGSGNRNAIIPVHDHYACDPRFRFILMPKNV-----	118				
QY	163	DGNYPQWPEAAGAVGAGAYREVEAEDPGCLAVEALVTRRCVCVCAQRMGGKREVMYTA	222				
DB	119	-----GKKRAEIVA	127				
QY	223	FKALGDSVDYVQVCDSDTRLDPMALLELVRVLDEDPRGVAGGVDVRIINPLDSDWVSFLSS	282				
DB	128	IRE--SSGDLVNLVDSDTIAFDVVVTKLAKM-YSFAVGAAGQLTASNRSDTLWTLFLID	184				
QY	283	LRVYVAFNVNVERACQSYFHCVSCISGSLGTPPGPAATQRRP-----RPSC	326				
DB	185	MEYWLACNEERRAAQSRFGAVMCCC-----GPCMYRRSALILLDKYETQLFRGRPSD	237				

Qy 327 GGASRH 332
Db 238 FGEDRH 243

RESULT 8
S12793
modulation protein nodC - Rhizobium loti
C;Species: Rhizobium loti
C;Date: 30-Sep-1993 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: S12793
R;Collins-Emerson, J.M.; Terzaghi, E.A.; Scott, D.B.
Nucleic Acids Res. 18, 6690, 1990
A;Title: Nucleotide sequence of Rhizobium loti nodC.
A;Reference number: S12793; MUID:91067465; PMID:2251130
A;Accession: S12793
A;Molecule type: DNA
A;Residues: 1-424 <COL>
A;Cross-references: UNIPROT:P17862; EMBL:X52958; NID:g46227; PIDN:CAA37131.1; PID:g46228
C;Comment: This is one of the proteins, coded by nodulation genes, that are required for
C;Genetics:
A;Gene: nodC
C;Superfamily: nodulation protein nodC
C;Keywords: nodulation

Query Match 8.9%; Score 176.5; DB 1; Length 424;
Best Local Similarity 26.6%; Pred. No. 2.3e-06;
Matches 83; Conservative 33; Mismatches 107; Indels 89; Gaps 13;

Qy 38 WYAGVPLASDRYGLLAGFLYGAFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSV 97
Db 4 FATASTVAICS--YALLS-----TVYKTAQVF--YTLPTNVAPTSGDPAGGEPPSV 51
Qy 98 ALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAEDLYMVMDFREYFADEDP 157
Db 52 DVIIPSYNEAPRTLSDCLASTASQEY-AGKLVVVVDGSG-----ANRDA 95
Qy 158 ATYWDGNYHQPWEPAAGAGVAGAYREVEAEDPGRLAVEALVTRRCVCAQRWGKGRE 217
Db 96 LVGVQEEYGHDP-----RNFIALPKN-----VGKRR 122
Qy 218 VMYTAFA-KALGSDYVYVQVCDSTRLDPMALLELRVRLDDEPRVGAGDVRILNPLDSW 276
Db 123 AQIAAVRRSCGDLVNV---DSDTILAPDVVTRLRALKM-QQAIGAMGQLAASNRGETW 178
Qy 277 VSFLLSLRYVAFNVERACQSYFHCVSCISGSLGTPPGPAA-----TQR- 320
Db 179 LTRLIDMEYWLACNEERAQAARFGAVMCCC-----GPCAMYRRSALLSVLDQYETQRF 231
Qy 321 RPRPSCGGASRH 332
Db 232 RGKPSDFGEDRH 243

RESULT 9
JC6079
chitin synthase (EC 2.4.1.16) chsD - Emericella nidulans
C;Species: Emericella nidulans, Aspergillus nidulans
C;Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 20-Jun-2000
C;Accession: JC6079
R;Motowama, T.; Fujiwara, M.; Kojima, N.; Horiuchi, H.; Ohta, A.; Takagi, M.
Mol. Gen. Genet. 251, 442-450, 1996
A;Title: The Aspergillus nidulans genes chsA and chsD encode chitin synthases which have
A;Reference number: JC6079; MUID:96285568; PMID:8709948
A;Accession: JC6079
A;Molecule type: DNA
A;Residues: 1-1086 <NOT>
A;Cross-references: DDBJ:D83246; NID:g1688025; PIDN:BAAL1866.1; PID:g1688026
C;Comment: This enzyme functions in conidia formation.
C;Genetics:
A;Gene: chsD
A;Introns: 800/2

C;Superfamily: chitin synthase chs4
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 8.2%; Score 163.5; DB 2; Length 1086;
Best Local Similarity 22.3%; Pred. No. 6.9e-05;
Matches 59; Conservative 46; Mismatches 111; Indels 49; Gaps 8;

Qy 88 PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDG---NRAEDLYM 144
Db 563 PYGFLAHMCL-VTCYSEGEIGRTTLDSTALTDPNSHKSIVVICDGIKGEPEFST 621
Qy 145 VDMFREYFAD-----ED---PATYVWDGNYHQPWEPAAGAGVAGAYREVEAEDPGRLA 195
Db 622 PDIVLRMRDPIIPPEVEAFPYAVATGSKRHNMAKVAGFYDYGHSIIPVEMQQRVP 681
Qy 196 VEALVTRRCVCAQRWG-----GKR-----EVMYTAFAKAL--- 226
Db 682 MMIIV---KCGTPAEATAAKPGNKRKDSQIILMSFLQKVMFDRMTELEYEMFNLHV 738
Qy 227 ----GSDVDYVQVCDSTRLDPMALLELRVRLDDEPRVGAGDVRILNPLDSVVSFLSS 282
Db 739 TGIPDPDFYEVVLMVDADTKVFPDLSLTHMISAMVKDPVGMGLCGETKIANTKDTDSWVTMIQV 798
Qy 283 LRYWVAFNVERACQSYFHCVSCISG 307
Db 799 FEYFVSHQSKAFESVEGGVTCCLPG 823

RESULT 10
JQ0396
nodulation protein nodC - Azorhizobium caulinodans
C;Species: Azorhizobium caulinodans
A;Note: host Sesbania rostrata
C;Date: 07-Sep-1990 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: JQ0396
R;Goethals, K.; Gao, M.; Tomekpe, K.; Van Montagu, M.; Holsters, M.
Mol. Gen. Genet. 219, 289-298, 1989
A;Title: Common nodABC genes in Nod locus 1 of Azorhizobium caulinodans: nucleotide sequ
A;Reference number: JQ0393; MUID:90136519; PMID:2615763
A;Accession: JQ0396
A;Molecule type: DNA
A;Residues: 1-395 <GOE>
A;Cross-references: UNIPROT:Q07755; GB:L18897; NID:g1293899; PIDN:AAB51164.1; PID:g31029
A;Experimental source: strain ORS571
C;Comment: This is one of the proteins, coded by nodulation genes, that are required for
C;Genetics:
A;Gene: nodC
C;Superfamily: nodulation protein nodC
C;Keywords: nodulation

Query Match 8.0%; Score 159.5; DB 1; Length 395;
Best Local Similarity 25.2%; Pred. No. 4.6e-05;
Matches 78; Conservative 32; Mismatches 101; Indels 99; Gaps 12;

Qy 49 DRYGLLAGFLYGAFLSAHLVAQSLFAYLEHRRVAAAARGPLDA--ATARSVALTISAYOE 106
Db 5 DVIGLLATAAYVTILASAYKVQ-----FINVSSVTDVAGLES DALPLTFR-VDVIVPTFNE 59
Qy 107 DPAYLRQCLASARALLYPRARLRVLMVVDGNRAE-----DLYMVMDFREYFADEDPAT 159
Db 60 NSSTLLECVASICAQDY-RGPITIVVVDGSGTNKTSFHAVCDKYASDE-RFIFVELD--- 114
Qy 160 YVWDGNYHQPWEPAAGAGVAGAYREVEAEDPGRLA-VEALVTRRCVCAQRWGKGRE 218
Db 115 -----QNKGTAAQMEAIRRT----- 129
Qy 219 MYTAFAKLGSDYVYVQVCDSTRLDPMALLELRVRLDDEPRVGAGDVRILNPLDSVVS 278
Db 130 -----DGLILNVDSDTVIDKDVVTKLASSM-RAPNVGVMQQLVAKNERSWLT 178
Qy 279 FLSSRLRYWVAFNVERACQSYFHCVSCISGSLGTPPGPAATQRRP----- 322
Db 179 RLIDMEYWLACNEERIAQSRFGSVMCCC-----GPCAMYRRSAITPLLAEBYHQTFLG 231


```
Qy 323 RPSGASRH 332
|||
Db 232 RPSNFGEDRH 241

RESULT 11
S61886
chitin synthase (EC 2.4.1.16) CHS4 - Neurospora crassa
N:Alternate names: chitin synthase class IV
C:Species: Neurospora crassa
C>Date: 23-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S61886
R:Bin, A.B.; Specht, C.A.; Robbins, P.W.; Yarden, O.
Mol. Gen. Genet. 250, 214-222, 1996
A:Title: chs-4, a class IV chitin synthase gene from Neurospora crassa.
A:Reference number: S61886; MUID:96188842; PMID:8628221
A:Accession: S61886
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1195 <DIN>
A:Cross-references: UNIPROT:Q01285; EMBL:U25097; NID:g793937; PIDN:AA03563.1; PID:g7939
C:Genetics:
A:Gene: chs-4
A:Introns: 1042/1
C:Superfamily: chitin synthase chs4
C:Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase; transmembrane

Query Match 8.0%; Score 159; DB 2; Length 1195;
Best Local Similarity 22.2%; Pred. No. 0.00018;
Matches 69; Conservative 51; Mismatches 127; Indels 64; Gaps 11;

Qy 46 LASDRYGLLFLGYG----AFLSAHLVAQSLFAYLEHRRVAAARGPLDAAATARSVALTI 101
|||
Db 606 LKSDAYSSSPADGPGAGIHEAVVQPPSDWM-----PGFPLAHTICH-V 653

Qy 102 SAYGEDPAYLRQCLASARALLYPRARLVLMVVDG---NRAEDLYVDMFREVFADE--- 155
|||
Db 654 TAYSEGEVGTTLTDSIAMTDYPNSHKLVILVCDDGIKGGHEHSTPDIILGMKDHITH 713

Qy 156 ----DPATY--VWDGNVHQHPPEAAGAGAGAYREVEAEDPGLAVEALVTRRCVCA 209
|||
Db 714 PDDVEPFSYAVATGSKRHNMAKVYTFDYGTNSAIFLEKQQRVPMVMV---KCGTPA 770

Qy 210 QRW-----GKKR-----EVMTAFKAL-----GDSVDYVQV---C 236
|||
Db 771 EASKSKPGRKRDQSIIILMSFLQKMFEDERMTELEYEFNGLWKITGISDFEIVLMV 830

Qy 237 DSDTRLDPMALLEVRVLDPRVAGVGDRILNPLDSWVSFSLSLRYWAFNVERACQ 296
|||
Db 831 DADTKVPFDSLTHMISAMVKDPEIMGLCGETKIANKRASWVSIAIQVFYFVSHLAKAFE 890

Qy 297 SYFHCVSCISG 307
|||
Db 891 SVFGVGTCLPG 901

RESULT 12
S45879
chitin synthase (EC 2.4.1.16) 3 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YBR023c; protein YBR0305
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S45879; S46554; S22776; A39639; S17247
R:Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45875
A:Accession: S45879
A:Molecule type: DNA
A:Residues: 1-1165 <GRI>
A:Cross-references: UNIPROT:P29465; EMBL:Z35892; NID:gs36229; PIDN:CAA84965.1; PID:gs362
A:Experimental source: strain S288C
R:Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.
```

```
Yeast 10(Suppl.A), S75-S80, 1994
A:Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II fro
ly identified genes and a homologue of the SCO1 gene.
A:Reference number: S46551; MUID:94378725; PMID:8091864
A:Accession: S46554
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1165 <SMI>
A:Cross-references: EMBL:X76078; NID:g498748; PIDN:CAA53680.1; PID:g498752
A:Experimental source: strain S288C
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
R:Bulawa, C.E.
Mol. Cell. Biol. 12, 1764-1776, 1992
A:Title: CSD2, CSD3, and CSD4, genes required for chitin synthesis in Saccharomyces cere
bium species and Xenopus laevis.
A:Reference number: S22776; MUID:92195323; PMID:1532231
A:Accession: S22776
A:Molecule type: DNA
A:Residues: 1-1162, 'L', 1164-1165 <BUL>
A:Cross-references: EMBL:M73697; NID:g172103; PIDN:AAA34844.1; PID:g172104
R:Valdivieso, M.H.; Mol, P.C.; Shaw, J.A.; Cabib, E.; Duran, A.
J. Cell Biol. 114, 101-109, 1991
A:Title: CALL, a gene required for activity of chitin synthase 3 in Saccharomyces cerevi
A:Reference number: A39639; MUID:91268144; PMID:2050737
A:Accession: A39639
A:Molecule type: DNA
A:Residues: 67-1165 <VAL>
A:Cross-references: GB:X57300; NID:g3359; PIDN:CAA40559.1; PID:g3360
C:Genetics:
A:Gene: SGD:CHS3; CALL; CSD2; MIPS:YBR023c
A:Cross-references: MIPS:YBR023c; SGD:S0000227
A:Map position: 2R
C:Function:
A:Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosa
C:Superfamily: chitin synthase chs4
C:Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase; transmembrane
F:203-279/Domain: transmembrane #status predicted <TM1>
F:457-473/Domain: transmembrane #status predicted <TM2>
F:1018-1034/Domain: transmembrane #status predicted <TM3>
F:1035-1054/Domain: transmembrane #status predicted <TM4>
F:1050-1076/Domain: transmembrane #status predicted <TM5>
F:1084-1100/Domain: transmembrane #status predicted <TM6>

Query Match 7.9%; Score 156.5; DB 1; Length 1165;
Best Local Similarity 22.2%; Pred. No. 0.00027;
Matches 58; Conservative 48; Mismatches 88; Indels 67; Gaps 9;

Qy 101 ISAYQEDPAYLRQCLASARALLYPRARLVLMVVDG---NRAED-----LYWVDMFRE 150
|||
Db 658 VTCYSEDEEGLRTTLDLSLTTPNSHKLLMVVCDGLIKSGNDKTTPEIAGLMMDDFVT 717

Qy 151 VFADEDPATYVWDGNVHQHPPEAAGA-----VGAGAYR-----EVEAEDPGLAVEAL 199
|||
Db 718 PPDEVKPYSYV-----AVASGSKRHNMAKIYAGFYKYDDSTIPPENQQRVPIITI 767

Qy 200 VTRRCVCVAQRMG-----GKR--EVMTAF-----KAL 226
|||
Db 768 V---KCGTPAEQGAAPGNRGRDQSIIILMSFLEKITFDERMTQLEFOLLKNIWITGLM 824

Qy 227 GDSVDYVQVCDSTRLDPMALLEVRVLDPRVAGVGDRILNPLDSWVSFSLSLRYW 286
|||
Db 825 ADFYETVLMVDADTKVPFDPALTHMVAEMVKDPLIMGLCGETKIANKAQSWVTAIQVFYY 884

Qy 287 VAFNVERACQSYFHCVSCISG 307
|||
Db 885 ISHQAKAFESVFGSVTCLPG 905

RESULT 13
T42020
class IV chitin synthase - smut fungus (Ustilago maydis)
C:Species: Ustilago maydis (corn smut)
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:46:32 ; Search time 105.75 Seconds

(without alignments)
1375.148 Million cell updates/sec

Title: US-10-672-399-4

Perfect score: 1987

Sequence: 1 MRQDAKPTPAARCSGLA.....TRNSPACSPSSWRPLCCVC 376

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	80.6	578	2	Aaw26765 Human hya
2	1509	75.9	582	3	Aay78135 Recombina
3	1509	75.9	583	2	Aaw30704 Mouse hya
4	1509	75.9	583	2	Aay32503 Hyaluroa
5	1509	75.9	583	3	Aay68491 Mouse hya
6	1509	75.9	583	3	Aab09948 Murine HA
7	1509	75.9	583	6	Abp96028 Mouse hya
8	1509	75.9	583	7	Aae39152 Mouse hya
9	1422	71.6	543	2	Aaw36503 Human hya
10	1416	71.3	582	3	Aay78138 Recombina
11	1416	71.3	583	3	Aay78131 Recombina
12	1412.5	71.1	584	3	Aay78132 Recombina
13	1235	62.2	563	3	Aay78129 Recombina
14	993.5	50.0	577	3	Aay78128 Recombina
15	989	49.8	573	3	Aay78127 Recombina
16	873.5	44.0	552	3	Aay78133 Recombina
17	873.5	44.0	553	3	Aay78136 Recombina
18	804	40.5	552	7	Adc49213 Rabbit hy
19	802.5	40.4	588	7	Add93928 Xenopus 1
20	793	39.9	554	2	Aaw50010 Murine hy
21	793	39.9	554	3	Aay68493 Mouse hya
22	793	39.9	554	6	Abp96030 Mouse hya
23	793	39.9	554	7	Aae39154 Mouse hya
24	793	39.9	554	8	Adj76214 Marker ge
25	787.5	39.6	552	3	Aay78140 Recombina

26	783.5	39.4	553	6	ABR48209 Human bla
27	783.5	39.4	553	6	ABU56490 Lung canc
28	783.5	39.4	553	6	ABU56686 Lung canc
29	783.5	39.4	553	7	ADN39052 Cancer/an
30	783.5	39.4	553	8	ADJ75425 Marker ge
31	783.5	39.4	553	8	ADN05116 Antipsoi
32	780.5	39.3	552	2	Aaw50009 Murine hy
33	780.5	39.3	552	3	Aay78141 Recombina
34	780.5	39.3	552	3	Aay78142 Recombina
35	780.5	39.3	552	3	Aay68492 Mouse hya
36	780.5	39.3	552	3	Aay78139 Recombina
37	780.5	39.3	552	5	ABb57265 Mouse isc
38	780.5	39.3	552	6	ABP96029 Mouse hya
39	780.5	39.3	552	7	ADA49684 Mouse hya
40	780.5	39.3	552	7	ADA49684 Human hya
41	780.5	39.3	552	7	Aae39153 Mouse hya
42	780.5	39.3	552	7	ADC59297 Human HAS
43	780.5	39.3	552	7	ADC59299 Mouse HAS
44	780.5	39.3	552	7	ADD48818 Human Pro
45	780.5	39.3	553	3	Aay78137 Recombina

ALIGNMENTS

RESULT 1

AAW26765
ID AAW26765 standard; protein; 578 AA.

XX AC AAW26765;

XX 21-MAY-1998 (first entry)

XX Human hyaluronan synthase.

XX Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
KW wound healing; vulnery; tissue repair; scar; keloid; therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Binding-site 119..127

FT Modified-site 340..343 /note= "putative hyaluronan binding site"

FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"

FT Modified-site 354..356 /note= "consensus phosphorylation sequence for protein kinase C"

FT Modified-site 371..374 /note= "consensus phosphorylation sequence for protein kinase C"

FT Modified-site 378..380 /note= "consensus phosphorylation sequence for cAMP-dependent protein kinase"

XX WO9740174-A1.

PD 30-OCT-1997.

XX 17-APR-1997; 97WO-US006350.

XX 22-APR-1996; 96US-00635552.

XX (LEUK-) LEUKOSITE INC.

XX Briekin MJ;

XX WPI; 1997-549359/50.

XX N-FSDB; AAT99541.

XX Human hyaluronan synthase - useful for recombinant production of
PT hyaluronic acid for wound healing, tissue repair and reducing

PT hypertrophic scar and keloid formation.

PS Claim 5; Page 38-40; 58pp; English.

XX This protein comprises human hyaluronan synthase (HAS), an enzyme
 CC involved in the synthesis of hyaluronan (hyaluronic acid) and which has
 CC the ability to confer cell adhesion by the lymphocyte receptor CD44. The
 CC amino acid sequence was deduced from clone 30C (see AAT99541), which was
 CC obtained from a human mesenteric lymph node expression library using an
 CC expression cloning system developed to isolate cDNA clones that encode
 CC proteins that confer adhesion of the murine T cell lymphoma TK1. The
 CC isolated clone can be utilised in a claimed method for producing HAS in
 CC host cells. Such host cells are used in a claimed method for the
 CC production of hyaluronan. Hyaluronan is useful for wound healing and
 CC tissue repair, and can reduce or prevent hypertrophic scars and keloid
 CC formation. It is also used in eye surgery as a replacement for vitreous
 CC fluid

SQ Sequence 578 AA;

Query Match 80.6%; Score 1602; DB 2; Length 578;
 Best Local Similarity 99.7%; Pred. No. 3.9e-154;
 Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
 DB 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
 QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCASARA 120
 DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCASARA 120
 QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREYFADEDPATYYWDCGNHYQHPWEPAAGAVGA 180
 DB 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREYFADEDPATYYWDCGNHYQHPWEPAAGAVGA 180
 QY 181 GAYREVEADDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
 DB 181 GAYREVEADDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
 QY 241 RLDPMALLELVRVLDDEPRVGAVGDDVRIINPLDSWVSFLSSRLRYWVAFNVERACQSYFH 300
 DB 241 RLDPMALLELVRVLDDEPRVGAVGDDVRIINPLDSWVSFLSSRLRYWVAFNVERACQSYFH 300
 QY 301 CVSCISGSLG 310
 DB 301 CVSCISGSLG 310

RESULT 2

AA78135
 ID AA78135 standard; protein; 582 AA.

AC AA78135;

DT 27-APR-2000 (first entry)

XX Recombinant chimeric hyaluronate synthase modified protein #9.

XX Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
 KW research reagent; biochemical research; medical development; chimeric.

XX Mus sp.

OS Synthetic.

OS Chimeric.

XX JP2000004886-A.

XX 11-JAN-2000.

XX 24-JUN-1998; 98JP-00193788.

XX 24-JUN-1998; 98JP-00193788.

XX (SEKG) SEIKAGAKU KOGYO CO LTD.

XX WPI; 2000-140125/13.

XX A hyaluronate synthase modified protein - useful as a research reagent
 PT for biochemical research and medical development.

XX Claim 10; Page; 30pp; Japanese.

XX The present invention describes a recombinant protein which consists of
 CC three continuous regions (N-terminal region, internal region and C-
 CC terminal region) where one or two regions among the above three regions
 CC is selected from the three hyaluronate synthase (HAS) modified proteins
 CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
 CC HAS2 amino acids 52 to 379 and HAS3 amino acids 410 to 583). The protein
 CC is useful as a research reagent for biochemical research and medical
 CC development. The invention provides a HAS modified protein of increased
 CC or lowered activity. AAY78127 to AAY78142 represent specifically claimed
 CC examples of recombinant proteins from the present invention. N.B. The
 CC present sequence is not given in the present specification, but is
 CC derived from sequences given as specified in the claim

XX Sequence 582 AA;

Query Match 75.9%; Score 1509; DB 3; Length 582;
 Best Local Similarity 93.9%; Pred. No. 1.3e-144;
 Matches 295; Conservative 2; Mismatches 11; Indels 6; Gaps 2;

QY 3 QDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLYCAF 62
 DB 3 QDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLYCAF 61
 QY 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCAS 117
 DB 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCAS 121
 QY 118 ARALLYPRLRLVLMVDGNRAEDLYMVDMPREYFADEDPATYYWDCGNHYQHPWEPA-AAG 176
 DB 122 ARALLYPRLRLVLMVDGNRAEDLYMVDMPREYFADEDPATYYWDCGNHYQHPWEPAATG 181
 QY 177 AVGAGAYREVEADDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVC 236
 DB 182 AVGAGAYREVEADDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVC 241
 QY 237 DSOTRLDPMALLELVRVLDDEPRVGAVGDDVRIINPLDSWVSFLSSRLRYWVAFNVERACQ 296
 DB 242 DSOTRLDPMALLELVRVLDDEPRVGAVGDDVRIINPLDSWVSFLSSRLRYWVAFNVERACQ 301
 QY 297 SYFHCVSCISGSLG 310
 DB 302 SYFHCVSCISGSLG 315

RESULT 3

AAW30704

ID AAW30704 standard; protein; 583 AA.

AC AAW30704;

DT 17-OCT-2003 (revised)

DT 19-DEC-1997 (first entry)

XX Mouse hyaluronate synthase enzyme.

DE Hyaluronate synthase; hyaluronic acid; polysaccharide.

XX Mus sp; (strain FM3A).

XX JP09224674-A.

XX 02-SEP-1997.

XX	Mus sp.
OS	
XX	WO2003006068-A1.
FN	
XX	23-JAN-2003.
PD	
XX	
PF	10-JUL-2001; 2001WO-US021785.
PR	
XX	10-JUL-2001; 2001WO-US021785.
PA	(CLEA-) CLEAR SOLUTIONS BIOTECH INC.
XX	
PI	Dehazya P, Chen W;
XX	
DR	WPI; 2003-221664/21.
DR	N-PSDB; ABZ76734.
XX	
PT	Novel dihydrazone derivatized hyaluronic acid/nucleic acid bioconjugate
PT	for treating dry eye syndrome, has derivatized hyaluronic acid
PT	crosslinked to nucleic acid encoding protein with hyaluronan synthase
PT	activity.
XX	
PS	Claim 20; Page 58-59; 62pp; English.
XX	
CC	The present invention describes a dihydrazone derivatised hyaluronic acid
CC	(HA)/nucleic acid (NA) bioconjugate (I), with derivatised HA linked to NA
CC	encoding protein with hyaluronan synthase (HAS) activity, where NA has
CC	sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659
CC	or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding
CC	proteins with 68-90% SI to a sequence of 583, 552 or 554 amino acids (see
CC	ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and
CC	antiarthritic activities, and can be used in gene therapy and as an
CC	inhibitor of angiogenesis, and as an inducer of expression of (HA) in
CC	human corneal epithelial cell. (I) is useful for transfecting a cell of
CC	an eye with nucleic acid, where the nucleic acid comprises (ABZ76734 to
CC	ABZ76736), by contacting the cell with (I) comprising the nucleic acid.
CC	(I) is useful for treating dry eye syndrome in an individual. (I) is
CC	useful in gene therapy applications for the treatment of a variety of
CC	medical conditions including dry eye syndrome or other medical conditions
CC	where an increase in the production of (HA) in the eye would be
CC	therapeutic (e.g., osteoarthritis of the articular joints). (I) is also
CC	useful for inhibiting angiogenesis for the treatment of macular
CC	degeneration or genes related to lipid biosynthesis that helps to restore
CC	the lipid component of the tear film, and as reagents for in vitro
CC	transformation of any cell, preferably a eukaryotic cell, more preferably
CC	a human eye cell. The present sequence represents mouse HAS1 which is
CC	used in the exemplification of the present invention
XX	
SQ	Sequence 583 AA;
	Query Match 75.9%; Score 1509; DB 6; Length 583;
	Best Local Similarity 93.9%; Pred. No. 1.3e-144;
	Matches 299; Conservative 2; Mismatches 11; Indels 6; Gaps 2
Qy	3 QODAPKETPAARRCSGLARRVLTATFALLILGLMTWAAGVPLASDRYGLLAFLGYCAF 62
	: : : : : : : : : : : :
Db	2 RQDMPKESERARCCSGLARALTIFALLILGLMTWAAGVPLASDRYGLLAFLGYCAF 61
Qy	63 LSAHLVAQSUFAYLEHRRVAAAAAR-----GPLDAATARSVALTTISAYQEDPAYLRQCILAS 117
Db	62 LSAHLVAQSUFAYLEHRRVAAAAARRSLAKGPLDAATARSVALTTISAYQEDPAYLRQCILTS 121
Qy	118 ARALLYPRRLRLVMVDGNRAEDLYMVDMEPREVFADEDPATYVWDGNYHQWPPEA-NAG 176
Db	122 ARALLYPHTRRLRLVMVDGNRAEDLYMVDMEPREVFADEDPATYVWDGNYHQWPPEAATG 181
Qy	177 AVGAGAYREVEAEPGRILAVALVRTRRCVCVAQRWGKGKRWMTAFKALGDSVDYVOVC 236
Db	182 AVGEGAYREVEAEPGRILAVALVRTRRCVCVAQRWGKGKRWMTAFKALGDSVDYVOVC 241
Qy	237 DSOTRDPMLLELRLVLDDEPPRGAVGGDVRLNPLDSWSFSLSLRYWYAFNVACQ 296

Db	242	DSDFRLDPMALLELVRVLDEDP	RVGAVGGDVRILNPLDSWVSFLSSLRVWVAFNVERACQ	30
Qy	297	SYFHCVSCISGSLG	310	
Db	302	SYFHCVSCISGPLG	315	
RESULT 8				
AAE39152				
ID	AAE39152	standard; protein; 583 AA.		
XX	AC	AAE39152;		
XX	XX	18-DEC-2003 (first entry)		
XX	XX	Mouse hyaluronan synthase (HAS) 1.		
XX	XX	Mouse; hyaluronic acid; HA; hyaluronan synthase; HAS; dry eye syndrome; eye cell; osteoarthritis; gene therapy; enzyme.		
XX	OS	Mus sp.		
XX	XX	US2003087850-A1.		
XX	XX	08-MAY-2003.		
XX	XX	10-JUL-2001; 2001US-00902939.		
XX	XX	10-JUL-2001; 2001US-00902939.		
XX	XX	(DEHA/) DEHAZYA P.		
XX	XX	(CHEN/) CHEN W.		
XX	XX	Dehazy P, Chen W;		
XX	XX	WPI; 2003-755151/71.		
XX	XX	N-PSDB; AAD59442.		
XX	XX	Dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate comprising derivatized hyaluronic acid cross linked to nucleic acid encoding hyaluronan synthase useful for treating dry eye syndrome.		
XX	XX	Claim 18; Page 15-16; 31pp; English.		
XX	XX	The invention relates to dihydrazide derivatised hyaluronic acid (HA) / nucleic acid bioconjugate comprising derivatised HA cross linked to nucleic acid encoding hyaluronan synthase (HAS). The invention is useful for treating an eye cell of an individual. It is useful for treating dry eye syndrome and osteoarthritis of the particular joints. The invention is also useful in gene therapy. The present sequence is mouse HAS1 enzyme		
XX	XX	Sequence 583 AA;		
XX	XX	Query Match 75.9%; Score 1509; DB 7; Length 583;		
XX	XX	Best Local Similarity 93.9%; Pred. No. 1.3e-144;		
XX	XX	Matches 295; Conservative 2; Mismatches 11; Indels 6; Gaps 2		
Qy	3	QQDAPKPTPAARCSGLARVLITAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAP	62	
Db	2	RQDMPKPEAARCCSGLARALATIFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAF	61	
Qy	63	LSAHLVAQSLPAYLEHRRVAAAAR----	GPLDAATARSVALTISAYOEDPAYLRQCLAS	117
Db	62	LSAHLVAQSLPAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYOEDPAYLRQCLTS	121	
Qy	118	ARALLYPRLRLVLMVVDGNRAEDLYMVDMFREFVFADEDPATYVMDGNYNHQPWEPA--AAG	176	
Db	122	ARALLYPHTRLRLVLMVVDGNRAEDLYMVDMFREFVFADEDPATYVMDGNYNHQPWEPAATG	181	
Qy	177	AVGAGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVC	236	
Db	182	AVGEGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVC	241	

Db 122 ABALYPPHTRLVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNYHQWPEAATG 181
QY 177 AVGAGAYREVEAEDPGRLAVALVTRTRCCVCAQWGGKREVMYTAFAKALGDSVDYVQC 236
Db 182 AVGAGAYREVEAEDPGRLAVALVTRTRCCVCAQWGGKREVMYTAFAKALGDSVDYVQC 241
QY 237 DSDTRLDPMALLEVRVLDDEPRVGAAGDVRILNPLDSWVSLSLRYWVAFNVERACQ 296
Db 242 DSDTMLDPASSVEMVKVLEEDPMVGGVGDVQILNKYDSWISFLSSVRYWMAFNTERACQ 301
QY 297 SYFHCVCISGSLG 310
Db 302 SYFGCVQCISGPLG 315

RESULT 11
AA78131
ID AAY78131 standard; protein; 583 AA.
XX AAY78131;
XX 27-APR-2000 (first entry)
DE Recombinant chimeric hyaluronate synthase modified protein #5.
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX JP2000004886-A.
XX 11-JAN-2000.
XX 24-JUN-1998; 98JP-00193788.
XX 24-JUN-1998; 98JP-00193788.
XX (SEKK) SEIKAGAKU KOGYO CO LTD.
XX WPI; 2000-140125/13.
XX A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
PS Claim 6; Page; 30pp; Japanese.

The present invention describes a recombinant protein which consists of three continuous regions (N-terminal region, internal region and C-terminal region) where one or two regions among the above three regions is selected from the three hyaluronate synthase (HAS) modified proteins HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71, HAS2 amino acids 52 to 379 and HAS3 amino acids 410 to 583). The protein is useful as a research reagent for biochemical research and medical development. The invention provides a HAS modified protein of increased or lowered activity. AAY78127 to AAY78142 represent specifically claimed examples of recombinant proteins from the present invention. N.B. The present sequence is not given in the present specification, but is derived from sequences given as specified in the claim

Sequence 583 AA;
Query Match 71.3%; Score 1416; DB 3; Length 583;
Best Local Similarity 87.3%; Pred. No. 4e-135;
Matches 274; Conservative 13; Mismatches 21; Indels 6; Gaps 2;
QY 3 QODAPKPTAARRCCGLARRVLTITAFALLILGLMTWAYAGVPLASDRYGLAFGLYGF 62
Db 2 RODMPKPEAARRCCGLARRVLTITAFALLILGLMTWAYAGVPLASDRYGLAFGLYGF 61

QY 63 LSAHLVAQSLFAYLEHRRVAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCCLAS 117
Db 62 LSAHLVAQSLFAYLEHRRVAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCCLTS 121
QY 118 ABALYPPHTRLVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNYHQWPEA- AAG 176
Db 122 ABALYPPHTRLVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNYHQWPEAATG 181
QY 177 AVGAGAYREVEAEDPGRLAVALVTRTRCCVCAQWGGKREVMYTAFAKALGDSVDYVQC 236
Db 182 AVGAGAYREVEAEDPGRLAVALVTRTRCCVCAQWGGKREVMYTAFAKALGDSVDYVQC 241
QY 237 DSDTRLDPMALLEVRVLDDEPRVGAAGDVRILNPLDSWVSLSLRYWVAFNVERACQ 296
Db 242 DSDTMLDPASSVEMVKVLEEDPMVGGVGDVQILNKYDSWISFLSSVRYWMAFNTERACQ 301
QY 297 SYFHCVCISGSLG 310
Db 302 SYFGCVQCISGPLG 315

RESULT 12
AA78132
ID AAY78132 standard; protein; 584 AA.
XX AAY78132;
XX 27-APR-2000 (first entry)
DE Recombinant chimeric hyaluronate synthase modified protein #6.
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX JP2000004886-A.
XX 11-JAN-2000.
XX 24-JUN-1998; 98JP-00193788.
XX 24-JUN-1998; 98JP-00193788.
XX (SEKK) SEIKAGAKU KOGYO CO LTD.
XX WPI; 2000-140125/13.
XX A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
PS Claim 7; Page; 30pp; Japanese.

The present invention describes a recombinant protein which consists of three continuous regions (N-terminal region, internal region and C-terminal region) where one or two regions among the above three regions is selected from the three hyaluronate synthase (HAS) modified proteins HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71, HAS2 amino acids 52 to 379 and HAS3 amino acids 410 to 583). The protein is useful as a research reagent for biochemical research and medical development. The invention provides a HAS modified protein of increased or lowered activity. AAY78127 to AAY78142 represent specifically claimed examples of recombinant proteins from the present invention. N.B. The present sequence is not given in the present specification, but is derived from sequences given as specified in the claim

Sequence 584 AA;
Query Match 71.1%; Score 1412.5; DB 3; Length 584;
Best Local Similarity 87.3%; Pred. No. 9.1e-135;
Matches 275; Conservative 13; Mismatches 20; Indels 7; Gaps 3;
SQ

[illegible]

RESULT 15	
AA78127	
ID	AA78127 standard; protein; 573 AA.
XX	
XX	AA78127;
XX	AC
XX	27-APR-2000 (first entry)
XX	
XX	Recombinant chimeric hyaluronate synthase modified protein #1.
XX	
DE	Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW	research reagent; biochemical research; medical development; chimeric.
KW	
XX	
OS	Mus sp.
OS	Synthetic.
OS	Chimeric.
XX	
XX	JP2000004886-A.
PN	
XX	
XX	11-JAN-2000.
XX	
XX	
PF	24-JUN-1998; 98JP-00193788.
XX	
XX	
PR	24-JUN-1998; 98JP-00193788.
XX	
XX	(SEKG) SEIKAGAKU KOGYO CO LTD.
PA	
XX	
XX	WPI; 2000-140125/13.
DR	
XX	
XX	
PT	A hyaluronate synthase modified protein - useful as a research reagent
PT	for biochemical research and medical development.
XX	
PS	Claim 2; Page; 30pp; Japanese.

CC	development.. The invention provides a HAS modified protein of increased or lowered activity.
CC	AAY78127 to AAY78142 represent specifically claimed examples of recombinant proteins from the present invention. N.B. The present sequence is not given in the present specification, but is derived from sequences given as specified in the claim
XX	CC
SQ	Sequence 573 AA;
	Query Match 49.8%; Score 989; DB 3; Length 573; Best Local Similarity 61.1%; Pred. No. 1.6e-91; Matches 195; Conservative 39; Mismatches 59; Indels 26; Gaps 5;
QY	3 QODAPKPTPAARCCSGLARVLITAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYCAF 62 : :
Dd	2 RQDMPKPSEAARCCSGLARRALTIIIFALLILGLMTWAYAAGVPLASDRYGLLAFGLYCAF 61 : :
QY	63 LSAHLVA-----QSFLPAYLEHRRVAAAARGPLDDATARSVALTISANYOEDPAYS 111 : :
Dd	62 LSAHLVAQSFLASHLIIQSFLAEFLEHRMKKSLETPI--KLNKTVALCIAANYOEDPDYL 119 : :
QY	112 ROCLASARALLYPRARLVLMVDGNAEDLYVMDFREVFADDEDPATYVMDGNHYHPWE 171 : :
Dd	120 RKCLQSVKKLTYP--GIKVMMVIDGNSDDLTYMNDIPSEVMGRDKSATIYWKNFHK-- 175 : :
QY	172 PAAAGAVGAGAREVEARDPPRLAVEALVRTRCVCVAQRGWGGKREVMYTAFKALGDSDV 231 : :
Dd	176 -----GFGETEESHKESSQH--VTQLVLNKNKSI CIMQKWGGKREVMYTAFRALGRSDV 226 : :
QY	232 YVQCDSDTLRIDPMALLELRVLDEDPRYGAVGGDVRIILNPDLSDWSVSFFLSRLVWAFNV 291 : :
Dd	227 YVQCDSDTMLDPASSVEMVKVLEEOPMGVGVDVQILNKYDSWISFLSSRVYWMAFNI 286 : :
QY	292 ERACQSYFHCVSCISGLG 310 : :
Dd	287 ERACQSYFCGCVICISGPLG 305 : :

Search completed: March 11, 2005, 14:17:13
Job time : 107.75 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:22:15 ; Search time 84.3864 Seconds
(without alignments)
1469.688 Million cell updates/sec

Title: US-10-672-399-4
Perfect score: 1987
Sequence: 1 MRQDAPKPTPAARRCSGLA.....TRNSPACSPSWRPLCCVC 376

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1987	100.0	376	16	US-10-672-399-4
2	1612	81.1	360	16	US-10-672-399-6
3	1602	80.6	578	13	US-10-042-523-2
4	1502	80.6	578	16	US-10-672-399-2
5	1509	75.9	583	10	US-09-902-939-4
6	1363	68.6	320	16	US-10-672-399-8
7	802.5	40.4	587	13	US-10-042-523-3
8	802.5	40.4	588	15	US-10-309-560-10
9	793	39.9	554	10	US-09-902-939-6
10	783.5	39.4	553	15	US-10-235-027-370
11	783.5	39.4	553	15	US-10-188-832-137
12	780.5	39.3	552	10	US-09-902-939-5
13	780.5	39.3	552	14	US-10-262-526-2
14	780.5	39.3	552	14	US-10-262-526-4
15	780.5	39.3	552	14	US-10-262-526-7
16	780.5	39.3	552	14	US-10-262-526-10
17	780.5	39.3	552	14	US-10-262-526-12
18	780.5	39.3	552	14	US-10-262-526-15
19	780.5	39.3	552	14	US-10-262-526-17
20	780.5	39.3	552	14	US-10-262-526-18
21	780.5	39.3	552	14	US-10-262-526-21
22	780.5	39.3	552	14	US-10-262-526-22
23	780.5	39.3	552	14	US-10-262-526-23
24	780.5	39.3	552	14	US-10-262-526-24
25	780.5	39.3	552	14	US-10-262-526-25
26	780.5	39.3	552	14	US-10-262-526-26
27	780.5	39.3	552	14	US-10-262-526-27
28	780.5	39.3	552	14	US-10-262-526-28
29	780.5	39.3	552	14	US-10-262-526-29
30	780.5	39.3	552	14	US-10-262-526-30
31	780.5	39.3	552	14	US-10-262-526-31
32	780.5	39.3	552	14	US-10-262-526-32
33	780.5	39.3	552	14	US-10-262-526-33
34	780.5	39.3	552	14	US-10-262-526-34
35	780.5	39.3	552	14	US-10-262-526-35
36	780.5	39.3	552	14	US-10-262-526-36
37	780.5	39.3	552	14	US-10-262-526-37
38	780.5	39.3	552	14	US-10-262-526-38
39	780.5	39.3	552	14	US-10-262-526-39
40	780.5	39.3	552	14	US-10-262-526-40
41	780.5	39.3	552	14	US-10-262-526-41
42	780.5	39.3	552	14	US-10-262-526-42
43	780.5	39.3	552	14	US-10-262-526-43
44	780.5	39.3	552	14	US-10-262-526-44
45	780.5	39.3	552	14	US-10-262-526-45

14	780.5	39.3	552	14	US-10-262-526-4	Sequence 4, Appli
15	308	15.5	567	9	US-09-879-959-7	Sequence 7, Appli
16	308	15.5	567	14	US-10-172-527-7	Sequence 7, Appli
17	308	15.5	567	15	US-10-309-560-12	Sequence 12, Appli
18	308	15.5	567	15	US-10-309-560-15	Sequence 10, Appli
19	308	15.5	568	14	US-10-011-768B-10	Sequence 10, Appli
20	167.5	8.4	417	15	US-10-309-560-19	Sequence 15, Appli
21	167.5	8.4	417	15	US-10-309-560-21	Sequence 19, Appli
22	167.5	8.4	417	15	US-10-309-560-20	Sequence 20, Appli
23	167.5	8.4	417	15	US-10-309-560-21	Sequence 21, Appli
24	167.5	8.4	417	15	US-10-309-560-26	Sequence 26, Appli
25	167.5	8.4	417	15	US-10-309-560-27	Sequence 27, Appli
26	167.5	8.4	417	15	US-10-309-560-28	Sequence 28, Appli
27	167.5	8.4	417	15	US-10-309-560-29	Sequence 29, Appli
28	166.5	8.4	417	9	US-09-879-959-2	Sequence 2, Appli
29	166.5	8.4	417	14	US-10-011-768B-2	Sequence 2, Appli
30	166.5	8.4	417	14	US-10-011-771B-2	Sequence 2, Appli
31	166.5	8.4	417	14	US-10-172-527-2	Sequence 2, Appli
32	166.5	8.4	417	14	US-10-326-185-2	Sequence 2, Appli
33	166.5	8.4	417	15	US-10-309-560-2	Sequence 2, Appli
34	166.5	8.4	417	15	US-10-309-560-16	Sequence 16, Appli
35	166.5	8.4	417	15	US-10-309-560-17	Sequence 17, Appli
36	166.5	8.4	417	15	US-10-309-560-18	Sequence 18, Appli
37	166.5	8.4	417	15	US-10-309-560-22	Sequence 22, Appli
38	166.5	8.4	417	15	US-10-309-560-23	Sequence 23, Appli
39	166.5	8.4	417	15	US-10-309-560-24	Sequence 24, Appli
40	166.5	8.4	417	15	US-10-309-560-25	Sequence 25, Appli
41	159	8.0	1219	15	US-10-369-493-3212	Sequence 3212, Ap
42	156.5	7.9	1165	9	US-09-801-368-76	Sequence 76, Appli
43	156.5	7.9	1165	15	US-10-369-493-1432	Sequence 1432, Ap
44	156	7.9	1160	15	US-10-369-493-12860	Sequence 12860, A
45	150.5	7.6	750	14	US-10-156-761-12649	Sequence 12649, A

ALIGNMENTS

RESULT 1
US-10-672-399-4
; Sequence 4, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-4

Query Match 100.0%; Score 1987; DB 16; Length 376;
Best Local Similarity 100.0%; Pred. No. 2.4e-169;
Matches 376; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL	60
Db	1	MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL	60
Qy	61	AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA	120
Db	61	AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA	120
Qy	121	LLYPRLRLVLMVDGNRAEDLYVMDMPREVFADDPATYVWDGNHYQHPWEPAAAGVGA	180
Db	121	LLYPRLRLVLMVDGNRAEDLYVMDMPREVFADDPATYVWDGNHYQHPWEPAAAGVGA	180

Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 301 CVSCISGSLGTPPGPAATQRRPRPSCGASRHAGPSRTSVSGCTTRSGGTGMRGPTRRW 360
Db |||||
Qy 301 CVSCISGSLGTPPGPAATQRRPRPSCGASRHAGPSRTSVSGCTTRSGGTGMRGPTRRW 360
Db |||||
Qy 361 SPACSPSWRPLCCVC 376
Db |||||
Qy 361 SPACSPSWRPLCCVC 376
Db |||||

RESULT 2

US-10-672-399-6
; Sequence 6, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-6

Query Match 81.1%; Score 1612; DB 16; Length 360;
Best Local Similarity 98.7%; Pred. No. 8.7e-136;
Matches 311; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
Db |||||
Qy 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
Db |||||
Qy 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db |||||
Qy 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db |||||
Qy 121 LLYPRARLRVLMVVDGNAEDLYVMDMFREVFADDPATYVMDGNHYQHPWEPAAAGAVGA 180
Db |||||
Qy 121 LLYPRARLRVLMVVDGNAEDLYVMDMFREVFADDPATYVMDGNHYQHPWEPAAAGAVGA 180
Db |||||
Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 301 CVSCISGSLGTPPGP 315
Db |||||
Qy 301 CVSCISGPLESCPGP 315
Db |||||

RESULT 3

US-10-672-399-2
; Sequence 2, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic

Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-042-523-2

Query Match 80.6%; Score 1602; DB 13; Length 578;
Best Local Similarity 99.7%; Pred. No. 1.3e-134;
Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
Db |||||
Qy 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
Db |||||
Qy 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db |||||
Qy 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db |||||
Qy 121 LLYPRARLRVLMVVDGNAEDLYVMDMFREVFADDPATYVMDGNHYQHPWEPAAAGAVGA 180
Db |||||
Qy 121 LLYPRARLRVLMVVDGNAEDLYVMDMFREVFADDPATYVMDGNHYQHPWEPAAAGAVGA 180
Db |||||
Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 301 CVSCISGSLG 310
Db |||||
Qy 301 CVSCISGPLG 310
Db |||||

RESULT 4

US-10-672-399-2
; Sequence 2, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics

```
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 2
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-2

Query Match      80.6%; Score 1602; DB 16; Length 578;
Best Local Similarity 99.7%; Pred. No. 1.3e-134;
Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYG 60
DB 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYG 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPEAAAGVGA 180
DB 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPEAAAGVGA 180
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240
DB 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240
QY 241 RLDPMALLELVRLVDEDPKRGVAGDVRILNPLDSWVSFLSLRYWVAFNVERACQSYFH 300
DB 241 RLDPMALLELVRLVDEDPKRGVAGDVRILNPLDSWVSFLSLRYWVAFNVERACQSYFH 300
QY 301 CVCSICSGSLG 310
DB 301 CVCSICSGPLG 310

RESULT 5
US-09-902-939-4
; Sequence 4, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Phillip DeHazy
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: mouse
US-09-902-939-4

Query Match      75.9%; Score 1509; DB 10; Length 583;
Best Local Similarity 93.9%; Pred. No. 2.7e-126;
Matches 295; Conservative 2; Mismatches 11; Indels 6; Gaps 2;

QY 3 QDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYGAP 62
DB 2 RDMKPKSEARCCSGLARRALTIIFALLILGLMTWAYAAGVPLASDRYGLLAFLGYGAF 61
QY 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 117
DB 62 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 121

; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 2
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-2

Query Match      80.6%; Score 1602; DB 16; Length 578;
Best Local Similarity 99.7%; Pred. No. 1.3e-134;
Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYG 60
DB 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYG 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPEAAAGVGA 180
DB 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPEAAAGVGA 180
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240
DB 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240
QY 241 RLDPMALLELVRLVDEDPKRGVAGDVRILNPLDSWVSFLSLRYWVAFNVERACQSYFH 300
DB 241 RLDPMALLELVRLVDEDPKRGVAGDVRILNPLDSWVSFLSLRYWVAFNVERACQSYFH 300
QY 301 CVCSICSGSLG 310
DB 301 CVCSICSGPLG 310

RESULT 6
US-10-672-399-8
; Sequence 8, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 8
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-8

Query Match      68.6%; Score 1363; DB 16; Length 320;
Best Local Similarity 83.4%; Pred. No. 1.5e-113;
Matches 267; Conservative 0; Mismatches 1; Indels 52; Gaps 1;

QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPEAAAGVGA 180
DB 61 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPEAAAGVGA 120
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240
DB 121 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 180
QY 241 RLDPMALLELVRLVDEDPKRGVAGDVRILNPLDSWVSFLSLRYWVAFNVERACQSYFH 300
DB 181 RLDPMALLELVRLVDEDPKRGVAGDVRILNPLDSWVSFLSLRYWVAFNVERACQSYFH 240
QY 301 CVCSICSGSLG 310
DB 241 CVCSICSGPLG 310

RESULT 7
US-10-042-523-3
; Sequence 3, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
```

```

; Acids and Uses Thereof
;
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-042-523-3

Query Match 40.4%; Score 802.5; DB 13; Length 587;
Best Local Similarity 53.4%; Pred. No. 5.3e-63;
Matches 156; Conservative 51; Mismatches 76; Indels 9; Gaps 3;

Qy 22 RVLTIAPALLILGLMTWAYAAGVPLASDRYGLLAGFLYGAFLSAHLVAQSLFAYLEHRRV 81
Db 28 RIYYSGVLLATITAAAYAEFQVLKHEAILFSLGLYGLAMLHLMMQSLFAFLEIRRV 87

Qy 82 AAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAED 141
Db 88 ---NKSELPSCFKKTVALTIAGYQENPEYLIKLESCKYVKYPKDKLKIILVIDGNTDD 144

Qy 142 LYVDMFREVFADEDPATYVMDGNYPWBPAAAGAVGAGAYREVEA---EDPGRLAVERA 198
Db 145 AYMMEFKDVFHGEDVGYVWKGNYHTVKPE---ETNKGSCPEVSKPLNEDEGINMVEE 201

Qy 199 LVTRTRCVCAQRWGGKREVMYTAFAKLGSDVYVQVCDSDTRIDPMALLELVRLVDEDP 258
Db 202 LVNRKRCVICMQQWGGKREVMYTAFAIGTSVDYVQVCDSDTKLDELATVEMVKVLESND 261

Qy 259 RVGAVGGDVRILNPLDSSVLSRLRYWVAFNVERACQSYFHCVCSISGSLG 310
Db 262 MYGAVGGDVRILNPFDSIFSMSSRLRYWMAFNVERACQSYFDCVCSISGPLG 313

RESULT 8
US-10-309-560-10
; Sequence 10, Application US/10309560
; Publication No. US20030235893A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: KUMARI, KSHAMA
; TITLE OF INVENTION: HYALURONAN SYNTHASES AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 35541.082
; CURRENT APPLICATION NUMBER: US/10/309,560

```

```

; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/336,105
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Xenopus laevis
; US-10-309-560-10

Query Match 40.4%; Score 802.5; DB 15; Length 588;
Best Local Similarity 53.4%; Pred. No. 5.4e-63;
Matches 156; Conservative 51; Mismatches 76; Indels 9; Gaps 3;

Qy 22 RVLTIAPALLILGLMTWAYAAGVPLASDRYGLLAGFLYGAFLSAHLVAQSLFAYLEHRRV 81
Db 28 RIYYSGVLLATITAAAYAEFQVLKHEAILFSLGLYGLAMLHLMMQSLFAFLEIRRV 87

Qy 82 AAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAED 141
Db 88 ---NKSELPSCFKKTVALTIAGYQENPEYLIKLESCKYVKYPKDKLKIILVIDGNTDD 144

Qy 142 LYVDMFREVFADEDPATYVMDGNYPWBPAAAGAVGAGAYREVEA---EDPGRLAVERA 198
Db 145 AYMMEFKDVFHGEDVGYVWKGNYHTVKPE---ETNKGSCPEVSKPLNEDEGINMVEE 201

Qy 199 LVTRTRCVCAQRWGGKREVMYTAFAKLGSDVYVQVCDSDTRIDPMALLELVRLVDEDP 258
Db 202 LVNRKRCVICMQQWGGKREVMYTAFAIGTSVDYVQVCDSDTKLDELATVEMVKVLESND 261

Qy 259 RVGAVGGDVRILNPLDSSVLSRLRYWVAFNVERACQSYFHCVCSISGSLG 310
Db 262 MYGAVGGDVRILNPFDSIFSMSSRLRYWMAFNVERACQSYFDCVCSISGPLG 313

RESULT 9
US-09-902-939-6
; Sequence 6, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; APPLICANT: Weiliam Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/0H020-US0
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 554
; TYPE: PRT
; ORGANISM: mouse
; US-09-902-939-6

Query Match 39.8%; Score 793; DB 10; Length 554;
Best Local Similarity 55.4%; Pred. No. 3.5e-62;
Matches 163; Conservative 37; Mismatches 76; Indels 18; Gaps 5;

Qy 21 RRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAGFLYGAFLSAHLVAQSLFAYLEHRR 80
Db 10 RVVGTSLFALVVLGGILAAVYGVQFTHTEKHYLSFGYLGLAILGHLHLLIQSLFAFLEHRR 69

Qy 81 VAAARGPLD---AATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGN 137
Db 70 MRRAGR-PLKLHCSQRSRSVALCIAAYQEDPEYLRKCLRSQAQRIAFP--NLKVVVMVVDGN 126

Qy 138 RAEDLYVMDMREVP-ADDEDPATYVMDGNYPWBPAAAGAVGAGAYREVEAEDPGR LAV 196
Db 127 RQEDTYMLDIFHEVLGTEQAGPFVWRNSNFHEAGEGSTEASLQEGMER-----V 175

Qy 197 EALVTRRCRCVCAQRWGGKREVMYTAFAKLGSDSYVYVQVCDSDTRIDPMALLELVRLVDE 256

```


Db 176 RAVWASTFSCIMQKGGKREVMYTAFAKALGNSVDYIQVCDSDTVDLPACTIEMLRVLEEDP 235
Qy 257 DPRVAGVGDRILNPLDSWSFLSSRLRYWAFNVERACQSYFHCVCISGSLG 310
Db 236 DPQVGGVDVQILNKYDSWISFLSSVRYWMAFNVERACQSYFCVCQISGSLG 289

RESULT 10
US-10-295-027-370
; Sequence 370, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 370
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-370

Query Match 39.4%; Score 783.5; DB 15; Length 553;
Best Local Similarity 55.1%; Pred. No. 2.5e-61;
Matches 161; Conservative 35; Mismatches 81; Indels 15; Gaps 4;
Qy 21 RRVLTAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRR 80
Db 10 RVVGTSLFALAVLGGILAAVYTGQFIHTEKHYSFGLYGAILLHLIQLSLFAFLEHRR 69
Qy 81 VAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRABRLVLMVVDGNRA 139
Db 70 MRRAGQALKLPSPRRGSVALCIAAYQEDPDYLRKCLRSARISFP--DLKVMVMVDGNRQ 127
Qy 140 EDLYMVDMFREVF-ADEDPATYVMDGNVHQPWEPAAAGAVGAGAYREVEADPGRLAYEA 198
Db 128 EDAYMLDIFHEVLGTEQAGFFVWRNPFHEAGEGTEASLQEGMDR-----VRD 176
Qy 199 LVRTRRCVCAQRWGGKREVMYTAFAKALGSDSYDVYQVCDSDTVDLPACTIEMLRVLEEDP 258
Db 177 VRASTFSCIMQKGGKREVMYTAFAKALGSDSYDVYQVCDSDTVDLPACTIEMLRVLEEDP 236
Qy 259 RVGAVGGDRILNPLDSWSFLSSRLRYWAFNVERACQSYFHCVCISGSLG 310
Db 237 QVGGVGGVDVQILNKYDSWISFLSSVRYWMAFNVERACQSYFCVCQISGSLG 288

RESULT 12
US-09-902-939-5
; Sequence 5, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHaza
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME

Db 177 VRASTFSCIMQKGGKREVMYTAFAKALGSDSYDVYQVCDSDTVDLPACTIEMLRVLEEDP 236
Qy 259 RVGAVGGDRILNPLDSWSFLSSRLRYWAFNVERACQSYFHCVCISGSLG 310
Db 237 QVGGVGGVDVQILNKYDSWISFLSSVRYWMAFNVERACQSYFCVCQISGSLG 288

RESULT 11
US-10-188-832-137
; Sequence 137, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 137
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-137

Query Match 39.4%; Score 783.5; DB 15; Length 553;
Best Local Similarity 55.1%; Pred. No. 2.5e-61;
Matches 161; Conservative 35; Mismatches 81; Indels 15; Gaps 4;
Qy 21 RRVLTAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRR 80
Db 10 RVVGTSLFALAVLGGILAAVYTGQFIHTEKHYSFGLYGAILLHLIQLSLFAFLEHRR 69
Qy 81 VAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRABRLVLMVVDGNRA 139
Db 70 MRRAGQALKLPSPRRGSVALCIAAYQEDPDYLRKCLRSARISFP--DLKVMVMVDGNRQ 127
Qy 140 EDLYMVDMFREVF-ADEDPATYVMDGNVHQPWEPAAAGAVGAGAYREVEADPGRLAYEA 198
Db 128 EDAYMLDIFHEVLGTEQAGFFVWRNPFHEAGEGTEASLQEGMDR-----VRD 176
Qy 199 LVRTRRCVCAQRWGGKREVMYTAFAKALGSDSYDVYQVCDSDTVDLPACTIEMLRVLEEDP 258
Db 177 VRASTFSCIMQKGGKREVMYTAFAKALGSDSYDVYQVCDSDTVDLPACTIEMLRVLEEDP 236
Qy 259 RVGAVGGDRILNPLDSWSFLSSRLRYWAFNVERACQSYFHCVCISGSLG 310
Db 237 QVGGVGGVDVQILNKYDSWISFLSSVRYWMAFNVERACQSYFCVCQISGSLG 288


```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 567.
; TYPE: PRT
; ORGANISM: Chlorella virus PBCV-1
US-09-879-959-7

Query Match      15.5%; Score 308; DB 9; Length 567;
Best Local Similarity 29.3%; Pred. No. 1e-18;
Matches 89; Conservative 57; Mismatches 108; Indels 50; Gaps 10;

Qy 19 LARRVLTIAFALLIIG-----LMTWYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSL 72
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
15 ITSNLIAVGGASLIILAPAITGYLHMNIALSTINGVSAYGIFVPGFF-----LAQVL 66
Qy 73 FAYLEHRR-----VAAAARGPLDAATARSVALTISAYQEDPAYLROCLASARALLYPRARL 128
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67 FSELNRKRLRKWISLRPKGWNDV-----RLAVIIAGYREDPYMFQKLESVRDSYGNV-A 121

Qy 129 RVLMYVDGNRAEDLYMDMFREVFADDEDPATYVMDGNTHQWPMPAAAGAVGAGAYREVEA 188
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 RLICVIDGDEDDMMAAVYKAIYND-----NIKKP-----EFVLCE 159

Qy 189 EDPGRLAVALVTRRCVCVAQRWCGCKREVMTAFK--ALGDSVDYVQVCDSDTRLDEMA 246
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
160 DDKEGERIDS--DFSREDICVLQPHRGKRECLYTGFLAKMDFSVNAVVLIDSDTVLEKDA 217

Qy 247 LLELVRVLEDEPRVGAAGVDVRIINPLDSWVSFLSRLRYVAFNVVERACQSYEHCVSCIS 306
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 ILEVYPLACDPEIQAVAGECKIWN-TDTLLSLLVANRYISAFCVERSAQSFRTVQCVG 276

Qy 307 GSLG 310
Db : : : :
277 GPLG 280
```

Search completed: March 11, 2005, 14:44:37
Job time : 85.3864 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:48:17 ; Search time 98.2727 seconds
(without alignments)
1959.259 Million cell updates/sec

Title: US-10-672-399-4

Perfect score: 1987

Sequence: 1 MRQDAPKPTPAARRCSGLA.....TRWSPACSSWRPLCCVC 376

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	80.6	578	1 HAS1_HUMAN	Q92839 homo sapien
2	1594	80.2	578	2 Q9NS49	Q9NS49 homo sapien
3	1585	79.8	577	2 O81YH3	O81YH3 homo sapien
4	1542.5	77.6	583	2 Q8S742	Q8S742 papio anubi
5	1509	75.9	583	1 HAS1_MOUSE	Q61647 mus musculus
6	1504	75.7	583	2 Q8CH93	Q8CH93 rattus norv
7	926.5	46.6	458	2 Q6T488	Q6T488 brachydanio
8	804	40.5	552	2 Q9SM29	Q9SM29 oryctolagus
9	803	40.4	557	2 Q6W9J2	Q6W9J2 xenopus lae
10	802.5	40.4	588	1 HAS1_XENLA	P13563 xenopus lae
11	793.5	39.9	553	2 Q7SR37	Q7SR37 sus scrofa
12	793	39.9	554	1 HAS3_MOUSE	Q08650 mus musculus
13	791	39.8	554	2 Q8CEB9	Q8CEB9 mus musculus
14	791	39.8	554	2 Q8CH92	Q8CH92 rattus norv
15	790.5	39.8	554	2 Q9DG40	Q9DG40 brachydanio
16	783.5	39.4	553	1 HAS3_HUMAN	Q00219 homo sapien
17	783.5	39.4	553	2 Q9GRV2	Q9GRV2 homo sapien
18	781.5	39.3	552	2 Q9SM16	Q9SM16 oryctolagus
19	780.5	39.3	552	1 HAS2_HUMAN	Q92819 homo sapien
20	780.5	39.3	552	1 HAS2_MOUSE	P70312 mus musculus
21	780.5	39.3	552	2 Q8HZJ3	Q8HZJ3 equus caball
22	779.5	39.2	552	2 Q8SQ70	Q8SQ70 sus scrofa
23	778.5	39.2	552	1 HAS2_BOVIN	O97711 bos taurus
24	775.5	39.0	552	1 HAS2_RAT	O35776 rattus norv
25	774	39.0	552	2 Q9DG41	Q9DG41 brachydanio
26	768	38.7	552	1 HAS2_CHICK	O57424 gallus gall
27	748.5	37.7	583	1 HAS3_XENLA	O57428 xenopus lae
28	747.5	37.6	583	2 Q6AZI0	Q6AZI0 xenopus lae
29	746.5	37.6	551	1 HAS2_XENLA	O57427 xenopus lae
30	592.5	29.8	281	2 Q8WTZ0	Q8WTZ0 homo sapien
31	560.5	28.2	131	2 Q8BPN0	Q8BPN0 mus musculus

32 534 26.9 393 2 Q811Y6 Q811Y6 rattus norv
33 344 17.3 156 2 O18792 O18792 papio anubi
34 308 15.5 568 2 O84419 O84419 paramescium
35 308 15.5 568 2 Q9WFS9 Q9WFS9 paramescium
36 308 15.5 568 2 Q9WFT0 Q9WFT0 paramescium
37 285 14.3 134 2 Q90489 Q90489 brachydanio
38 265 13.3 245 2 Q9GK14 Q9GK14 bos taurus
39 213.5 10.7 452 2 Q9AQ23 Q9AQ23 brachyrihob
40 213 10.7 424 1 NODC_RHILV P04340 rhizobium l
41 208.5 10.5 426 1 NODC_RHIME P04341 rhizobium m
42 205.5 10.3 452 2 Q8EXE1 Q8EXE1 sinorhizobi
43 201.5 10.1 413 1 NODC_RHISN P50357 rhizobium s
44 200.5 10.1 408 2 Q6PTX8 Q6PTX8 rhizobium s
45 196.5 9.9 424 2 Q9R614 Q9R614 rhizobium m

ALIGNMENTS

RESULT 1

HAS1_HUMAN STANDARD; PRT; 578 AA.
AC Q92839; Q14470;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
GN Name=HAS1; Synonyms=HAS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RX MEDLINE=96394438; PubMed=8798544; DOI=10.1074/jbc.271.38.23395;
RA Shvjan A.M., Heldin P., Butcher E.C., Yoshino T., Briskin M.J.;
RT "Functional cloning of the cDNA for a human hyaluronan-synthase";
RJ J Biol Chem. 271:23395-23399(1996)
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96244584; PubMed=8651928; DOI=10.1006/bbrc.1996.0827;
RA Itano N., Kinata K.;
RT "Molecular cloning of human hyaluronan synthase";
RL Biochem Biophys Res Commun. 222:816-820(1996).
CC -- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA) synthesis. Also able to catalyze the synthesis of chito-oligosaccharide depending on the substrate.
CC -- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-glucuronosyl(1->3)](n) + 2n UDP.
CC -- COFACTOR: Magnesium.
CC -- PATHWAY: Hyaluronate synthesis.
CC -- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -- TISSUE SPECIFICITY: Highly expressed in ovary followed by spleen, thymus, prostate, testes and large intestine. Weakly expressed in small intestine.
CC -- SIMILARITY: Belongs to the nodC/HAS family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U59269; AAC50706.1; --
EMBL; D84424; BAA12351.1; ALT_INIT.
Genew; HGNC:4818; HAS1.
MIM; 601463; --

```

DR GO: 0005887; C: integral to plasma membrane; TAS.
DR GO: 0007155; P: cell adhesion; TAS.
DR GO: 0006024; P: glycosaminoglycan biosynthesis; TAS.
DR InterPro: IPR001173; Glyco_transf_2.
DR Pfam: PF00535; Glyco_transf_2; 1.
KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
FT DOMAIN 1 25 Cytoplasmic (Potential).
FT TRANSMEM 26 46 1 (Potential).
FT DOMAIN 47 52 Extracellular (Potential).
FT TRANSMEM 53 73 2 (Potential).
FT DOMAIN 74 399 Cytoplasmic (Potential).
FT TRANSMEM 400 420 3 (Potential).
FT DOMAIN 421 430 Extracellular (Potential).
FT TRANSMEM 431 451 4 (Potential).
FT DOMAIN 452 457 Cytoplasmic (Potential).
FT TRANSMEM 458 478 5 (Potential).
FT DOMAIN 479 497 Extracellular (Potential).
FT TRANSMEM 498 518 6 (Potential).
FT DOMAIN 519 540 Cytoplasmic (Potential).
FT TRANSMEM 541 561 7 (Potential).
FT DOMAIN 562 578 Extracellular (Potential).
FT CONFLICT 1 2 MR -> RS (in Ref. 2).
FT CONFLICT 34 34 G -> A (in Ref. 2).
SQ SEQUENCE 578 AA; 64884 MW; 355FD54B0899E43C CRC64;

Query Match 80.6%; Score 1602; DB 1; Length 578;
Best Local Similarity 99.7%; Pred. No. 1e-116;
Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
Db 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60

Qy 61 AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db 61 AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120

Qy 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYWDGNYHQWPWPAAGAUGA 180
Db 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYWDGNYHQWPWPAAGAUGA 180

Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240

Qy 241 RLDPMALLELRVLDDEPRVGAAGDVRILNPLDSWVSFLSSLRYYWAFNVERACQSYFH 300
Db 241 RLDPMALLELRVLDDEPRVGAAGDVRILNPLDSWVSFLSSLRYYWAFNVERACQSYFH 300

Qy 301 CVSCISGSLG 310
Db 301 CVSCISGSLG 310

RESULT 2
Q9NS49 PRELIMINARY; PRT; 578 AA.
AC Q9NS49;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hyaluronan synthase 1.
GN Name=HAS1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E.; McCready P.M.; Skowronski E.; Viswanathan V.;
RA Burkhardt-Schultz K.; Gordon L.; Dias J.; Ramirez M.; Stillwagen S.;
RA Phan H.; Velasco N.; Do L.; Regala W.; Terry A.; Brower A.; Ganes J.;
RA Danganan L.; Erler A.; Christensen M.; Georgescu A.; Avila J.; Liu S.;

RA Andreise T., Trankheim M., Attix C., Amico-Keller G., Cosfield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA DOE Joint Genome Institute;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E.;
RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC018755; AAF87845.1; -.
DR InterPro: IPR001173; Glyco_transf_2.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR002057; Isopen N synth.
DR PROSITE; PS00237; G PROTEIN RECEPTOR F1.1; UNKNOWN 1.
DR PROSITE; PS00185; IPNS 1; UNKNOWN 1.
DR PROSITE; PS00185; IPNS 1; UNKNOWN 1.
SQ SEQUENCE 578 AA; 64831 MW; 2FE3A44B0D5380FF CRC64;

Query Match 80.2%; Score 1594; DB 2; Length 578;
Best Local Similarity 99.4%; Pred. No. 4.2e-116;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
Db 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60

Qy 61 AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db 61 AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120

Qy 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYWDGNYHQWPWPAAGAUGA 180
Db 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYWDGNYHQWPWPAAGAUGA 180

Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240

Qy 241 RLDPMALLELRVLDDEPRVGAAGDVRILNPLDSWVSFLSSLRYYWAFNVERACQSYFH 300
Db 241 RLDPMALLELRVLDDEPRVGAAGDVRILNPLDSWVSFLSSLRYYWAFNVERACQSYFH 300

Qy 301 CVSCISGSLG 310
Db 301 CVSCISGSLG 310

RESULT 3
Q81YH3 PRELIMINARY; PRT; 577 AA.
AC Q81YH3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hyaluronan synthase 1.
GN Name=HAS1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Krausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Heish F.;
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;

```

RL	Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY463695; NAR25554.1; -
DR	InterPro; IPR001173; Glyco trans 2.
DR	InterPro; IPR000276; GPCR Rhodpn.
DR	InterPro; IPR002057; Isopen N synth.
DR	PROSITE; PS00237; G_PROTEIN_RECSP_F1_1; UNKNOWN_1.
DR	PROSITE; PS00185; IPNS 1; UNKNOWN_1.
SQ	SEQUENCE 583 AA; 65596 MW; EA47354C89346P94 CRC64;
Query Match	
Best Local Similarity 94.9%; Score 1542.5; DB 2; Length 583;	
Matches 299; Conservative 3; Mismatches 8; Indels 5; Gaps 1;	
QY	1 MRQODAPKPTPAARRCSGLARRVLTITAFALLIIGLMTWAYAAGVPLASDRYGLLAFGLYG 60
DB	1 MTQDTPKPTPAARRCSGLARRVLTITAFALLIIGLMTWAYAAGVPLASDRYGLLAFGLYG 60
QY	61 AFISAHILVAQSLPAYLEHRRVA-----AAARGPLDAATARSVALTISAYQEDPAYLRQCL 115
DB	61 AFISAHLLAQSLPAYLEHRRVAAAARAAARGPLDAATARSVALTISAYQEDPAYLRQCL 120
QY	116 ASARALLYPRARLURLVLMVDGNRAEDLYMVDMPREVFADDPATYYWDGNYHQPWEPAVA 175
DB	121 VSARALLYPRARLURLVLMVDGNRPEDLYMVDMPREVFADDPATYYWDGNYHQPWEPAV 180
QY	176 GAVGAGAYREVEAEDPGRLAVEALVLRTRCVCVAQRWGKGKREVMYTAFAKALGDSVDYVQV 235
DB	181 GAVGVGAYREVEAEDPGRLAVEALVLRTRCVCVAQRWGKGKREVMYTAFAKALGDSVDYVQV 240
QY	236 CDSDTLDPMLLELVRVLDEDDPRGAVGVDVRIILNPLDSWVSGFLSSRLRYWVAFNVERAC 295
DB	241 CDSDTLDPMLLELVLQVLDEDDPRGAVGVDVRIILNPLDSWVSGFLSSRLRYWVAFNVERAC 300
QY	296 QSYFHCVCISGSLG 310
DB	301 QSYFHCVCISGSLG 315
RESULT 5	
ID	HAS1_MOUSE STANDARD; PRT; 583 AA.
AC	Q61647;
DT	30-MAY-2000 (Rel. 39, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DE	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE	(Hyaluronic acid synthase 1) (HA synthase 1).
GN	Name=Has1; Synonyms=Has;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RN	SEQUENCE FROM N.A.
RP	MEDLINE=96215261; PubMed=8626618; DOI=10.1074/jbc.271.17.9875;
RA	Itano N., Kimata K.;
RA	Itano N., Kimata K.;
RT	"Expression cloning and molecular characterization of HAS protein, a
RT	eukaryotic hyaluronan synthase";
RT	for the activity.";
RL	J. Biol. Chem. 271:9875-9878(1996).
RN	[2]
RN	MUTAGENESIS.
RP	MEDLINE=20085071; PubMed=10617644; DOI=10.1074/jbc.275.1.497;
RA	Yoshida M., Itano N., Yamada Y., Kimata K.;
RA	"In vitro synthesis of hyaluronan by a single protein derived from
RT	mouse HAS1 gene and characterization of amino acid residues essential
RT	for the activity.";
RL	J. Biol. Chem. 275:497-506(2000).
CC	-I- FUNCTION: plays a role in hyaluronan/hyaluronic acid (HA)
CC	synthesis. Also able to catalyze the synthesis of chito-
CC	oligosaccharide depending on the substrate.
CC	-I- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
CC	glucuronate = [beta-N-acetyl-D-glucosaminyl (1->4)beta-D-
CC	glucuronosyl (1->3)] (n) + 2n UDP.

Db	10	RVVGTSLFALVAGGLAAAYVTGQYTHTEKHVLSFGLYGAILGLHLLIQLSFLAFLEHRR	69			
QY	81	VAAAAAGPLDAAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLVMVVDGNRAE	140			
Db	70	MRZARPLKLPSPRRSVALCIAAYQEDPDYLRKLSAQRIAFP--DLKVMVVDGNRQE	127			
QY	141	DLVMVDMFEV--ADEDPATYVDGNYHQWPPEAAGAGAGAYREVEAEDPGRLAVEAL	199			
Db	128	DAYMLDIFHEVLGGTEQAGPFVVRNSPHEAGEGETASLBQEMER-----VRV	176			
QY	200	VRTRRCCVQAORGGGKREVMYTAFAKLGDSVDYVQVCDSDTRLDPMALLLRVLVLDPR	259			
Db	177	VRTSTSCINQKGGGKREVMYTAFAKLGDSVDYVQVCDSDTRLDPMALLLRVLVLDPR	236			
QY	260	VGAVGGDVRIINPLDWSVFLSLRLRYVAFNVERACQSYFHCVSCISGSLG	310			
Db	237	VGGVGGDVQILNKYDWSVFLSLRVYVMAFNVERACQSYFGCVQICSGPLG	287			
RESULT 9						
Q6W9J2						
ID	Q6W9J2	PRELIMINARY;	PRT; 557 AA.			
AC	Q6W9J2;					
DT	05-JUL-2004	(T-EMBLrel. 27, Created)				
DT	05-JUL-2004	(T-EMBLrel. 27, Last sequence update)				
DT	05-JUL-2004	(T-EMBLrel. 27, Last annotation update)				
DE	Hyaluronic acid synthase 3.					
OS	Xenopus laevis (African clawed frog).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;					
OC	Xenopodinae; Xenopus.					
OX	NCBI_TaxID=8355;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	PubMed=14667843; DOI=10.1016/S0945-053X(03)00082-9;					
RA	Vigetti D., Viola M., Gornati R., Ori M., Nardi I., Passi A.,					
RA	De Luca G., Bernardini G.					
RT	"Molecular cloning, genomic organization and developmental expression					
RL	of the Xenopus laevis hyaluronan synthase 3."					
RL	Matrix Biol. 22:511-517(2003).					
DR	EMBL; AY302252; AAP58398.1; -					
DR	InterPro; IPR001173; Glyco trans 2.					
DR	Pfam; PF00535; Glycos transf 2; 1.					
SQ	SEQUENCE 557 AA; 64060 MW; F81478B685B9F7AD5 CRC64;					
Query Match						
Best Local Similarity 55.4%; Score 803; DB 2; Length 557;						
Matches 164; Conservative 36; Mismatches 71; Indels 24; Gaps 6						
QY	17	SGLARRVLTAFALLIILGLMTWYAAAGVP--LASDRYGLLAFGLYGAFLSAHLVAQSLFAY	75			
Db	7	TGL--RVLATCLFALLVLGGFLVAYVTGQYTHTRHH--LSFGLYGAILGLHLLSQSLPAF	64			
QY	76	LEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLVMVD	135			
Db	65	LEHRKMRGGGRCFSGKST---VVLCTAAQEDPEYLRKLSVRRLSYP--HLRVIMVD	119			
QY	136	GNRAEDLYVDMFREFVADDPATYVDGNYHQWPPEAAGAGAGAYREVEAEDPGRLA	195			
Db	120	GNTEEDRYMMDI PREVMGSEGTCCYIWDKNYHSEEGQEGERG-----	163			
QY	196	VEALVTRRCVCVQAORGGGKREVMYTAFAKLGDSVDYVQVCDSDTRLDPMALLLRVLVD	255			
Db	164	QVEKVNKFQYVCINQKGGKREVMYTAFAKLGDSVAIVQVCDSDTDLDPACTAEMRLILE	223			
QY	256	EDPRGAVGDNRIINPLDSVFLSLRLRYVAFNVERACQSYFHCVSCISGSLG	310			
Db	224	EDPEVGGVGDVQILNKYDWSVFLSLRVYVMAFNVERACQSYFGCVQICSGPLG	278			

ID HAS1_XENLA STANDARD; PRT; 588 AA.
 AC P13563;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
 DE (Hyaluronic acid synthase 1) (HA synthase 1) (XHAS1) (DG42 protein).
 GN Name=HAS1;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88313363; PubMed=3410156;
 RA Rosa F., Sargent T.D., Rebert M.L., Michaels G.S., Jamrich M.,
 RA Grunz H., Jonas E., Winkles J.A., Dawid I.B.;
 RT "Accumulation and decay of DG42 gene products follow a gradient
 RT pattern during Xenopus embryogenesis.";
 RL Dev. Biol. 129:114-123(1988).
 CC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
 CC synthesis (By similarity). May play a role in signaling or pattern
 CC formation in embryonic development.
 CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
 CC glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-
 CC glucuronosyl(1->3)](n) + 2n UDP.
 CC -!- COFACTOR: Magnesium.
 CC -!- PATHWAY: Hyaluronate synthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- TISSUE SPECIFICITY: Embryo.
 CC -!- DEVELOPMENTAL STAGE: During gastrula and early neurula stages.
 CC -!- SIMILARITY: Belongs to the nodC/HAS family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; M22249; AAA49699.1; --
 DR PIR; A43740; A43740.
 DR InterPro; IPR001173; Glyco_trans.2.
 KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
 FT DOMAIN 1 28
 FT TRANSFEM 29 49
 FT DOMAIN 50 61
 FT TRANSFEM 62 82
 FT TRANSFEM 83 411
 FT DOMAIN 412 432
 FT TRANSFEM 433 433
 FT TRANSFEM 434 454
 FT TRANSFEM 455 456
 FT TRANSFEM 457 477
 FT TRANSFEM 478 505
 FT TRANSFEM 506 526
 FT TRANSFEM 527 543
 FT TRANSFEM 544 564
 FT TRANSFEM 565 588
 FT TRANSFEM 589 600
 FT TRANSFEM 601 620
 FT TRANSFEM 621 640
 FT TRANSFEM 641 660
 FT TRANSFEM 661 680
 FT TRANSFEM 681 700
 FT TRANSFEM 701 720
 FT TRANSFEM 721 740
 FT TRANSFEM 741 760
 FT TRANSFEM 761 780
 FT TRANSFEM 781 800
 FT TRANSFEM 801 820
 FT TRANSFEM 821 840
 FT TRANSFEM 841 860
 FT TRANSFEM 861 880
 FT TRANSFEM 881 900
 FT TRANSFEM 901 920
 FT TRANSFEM 921 940
 FT TRANSFEM 941 960
 FT TRANSFEM 961 980
 FT TRANSFEM 981 1000
 FT TRANSFEM 1001 1020
 FT TRANSFEM 1021 1040
 FT TRANSFEM 1041 1060
 FT TRANSFEM 1061 1080
 FT TRANSFEM 1081 1100
 FT TRANSFEM 1101 1120
 FT TRANSFEM 1121 1140
 FT TRANSFEM 1141 1160
 FT TRANSFEM 1161 1180
 FT TRANSFEM 1181 1200
 FT TRANSFEM 1201 1220
 FT TRANSFEM 1221 1240
 FT TRANSFEM 1241 1260
 FT TRANSFEM 1261 1280
 FT TRANSFEM 1281 1300
 FT TRANSFEM 1301 1320
 FT TRANSFEM 1321 1340
 FT TRANSFEM 1341 1360
 FT TRANSFEM 1361 1380
 FT TRANSFEM 1381 1400
 FT TRANSFEM 1401 1420
 FT TRANSFEM 1421 1440
 FT TRANSFEM 1441 1460
 FT TRANSFEM 1461 1480
 FT TRANSFEM 1481 1500
 FT TRANSFEM 1501 1520
 FT TRANSFEM 1521 1540
 FT TRANSFEM 1541 1560
 FT TRANSFEM 1561 1580
 FT TRANSFEM 1581 1600
 FT TRANSFEM 1601 1620
 FT TRANSFEM 1621 1640
 FT TRANSFEM 1641 1660
 FT TRANSFEM 1661 1680
 FT TRANSFEM 1681 1700
 FT TRANSFEM 1701 1720
 FT TRANSFEM 1721 1740
 FT TRANSFEM 1741 1760
 FT TRANSFEM 1761 1780
 FT TRANSFEM 1781 1800
 FT TRANSFEM 1801 1820
 FT TRANSFEM 1821 1840
 FT TRANSFEM 1841 1860
 FT TRANSFEM 1861 1880
 FT TRANSFEM 1881 1900
 FT TRANSFEM 1901 1920
 FT TRANSFEM 1921 1940
 FT TRANSFEM 1941 1960
 FT TRANSFEM 1961 1980
 FT TRANSFEM 1981 2000
 FT TRANSFEM 2001 2020
 FT TRANSFEM 2021 2040
 FT TRANSFEM 2041 2060
 FT TRANSFEM 2061 2080
 FT TRANSFEM 2081 2100
 FT TRANSFEM 2101 2120
 FT TRANSFEM 2121 2140
 FT TRANSFEM 2141 2160
 FT TRANSFEM 2161 2180
 FT TRANSFEM 2181 2200
 FT TRANSFEM 2201 2220
 FT TRANSFEM 2221 2240
 FT TRANSFEM 2241 2260
 FT TRANSFEM 2261 2280
 FT TRANSFEM 2281 2300
 FT TRANSFEM 2301 2320
 FT TRANSFEM 2321 2340
 FT TRANSFEM 2341 2360
 FT TRANSFEM 2361 2380
 FT TRANSFEM 2381 2400
 FT TRANSFEM 2401 2420
 FT TRANSFEM 2421 2440
 FT TRANSFEM 2441 2460
 FT TRANSFEM 2461 2480
 FT TRANSFEM 2481 2500
 FT TRANSFEM 2501 2520
 FT TRANSFEM 2521 2540
 FT TRANSFEM 2541 2560
 FT TRANSFEM 2561 2580
 FT TRANSFEM 2581 2600
 FT TRANSFEM 2601 2620
 FT TRANSFEM 2621 2640
 FT TRANSFEM 2641 2660
 FT TRANSFEM 2661 2680
 FT TRANSFEM 2681 2700
 FT TRANSFEM 2701 2720
 FT TRANSFEM 2721 2740
 FT TRANSFEM 2741 2760
 FT TRANSFEM 2761 2780
 FT TRANSFEM 2781 2800
 FT TRANSFEM 2801 2820
 FT TRANSFEM 2821 2840
 FT TRANSFEM 2841 2860
 FT TRANSFEM 2861 2880
 FT TRANSFEM 2881 2900
 FT TRANSFEM 2901 2920
 FT TRANSFEM 2921 2940
 FT TRANSFEM 2941 2960
 FT TRANSFEM 2961 2980
 FT TRANSFEM 2981 3000
 FT TRANSFEM 3001 3020
 FT TRANSFEM 3021 3040
 FT TRANSFEM 3041 3060
 FT TRANSFEM 3061 3080
 FT TRANSFEM 3081 3100
 FT TRANSFEM 3101 3120
 FT TRANSFEM 3121 3140
 FT TRANSFEM 3141 3160
 FT TRANSFEM 3161 3180
 FT TRANSFEM 3181 3200
 FT TRANSFEM 3201 3220
 FT TRANSFEM 3221 3240
 FT TRANSFEM 3241 3260
 FT TRANSFEM 3261 3280
 FT TRANSFEM 3281 3300
 FT TRANSFEM 3301 3320
 FT TRANSFEM 3321 3340
 FT TRANSFEM 3341 3360
 FT TRANSFEM 3361 3380
 FT TRANSFEM 3381 3400
 FT TRANSFEM 3401 3420
 FT TRANSFEM 3421 3440
 FT TRANSFEM 3441 3460
 FT TRANSFEM 3461 3480
 FT TRANSFEM 3481 3500
 FT TRANSFEM 3501 3520
 FT TRANSFEM 3521 3540
 FT TRANSFEM 3541 3560
 FT TRANSFEM 3561 3580
 FT TRANSFEM 3581 3600
 FT TRANSFEM 3601 3620
 FT TRANSFEM 3621 3640
 FT TRANSFEM 3641 3660
 FT TRANSFEM 3661 3680
 FT TRANSFEM 3681 3700
 FT TRANSFEM 3701 3720
 FT TRANSFEM 3721 3740
 FT TRANSFEM 3741 3760
 FT TRANSFEM 3761 3780
 FT TRANSFEM 3781 3800
 FT TRANSFEM 3801 3820
 FT TRANSFEM 3821 3840
 FT TRANSFEM 3841 3860
 FT TRANSFEM 3861 3880
 FT TRANSFEM 3881 3900
 FT TRANSFEM 3901 3920
 FT TRANSFEM 3921 3940
 FT TRANSFEM 3941 3960
 FT TRANSFEM 3961 3980
 FT TRANSFEM 3981 4000
 FT TRANSFEM 4001 4020
 FT TRANSFEM 4021 4040
 FT TRANSFEM 4041 4060
 FT TRANSFEM 4061 4080
 FT TRANSFEM 4081 4100
 FT TRANSFEM 4101 4120
 FT TRANSFEM 4121 4140
 FT TRANSFEM 4141 4160
 FT TRANSFEM 4161 4180
 FT TRANSFEM 4181 4200
 FT TRANSFEM 4201 4220
 FT TRANSFEM 4221 4240
 FT TRANSFEM 4241 4260
 FT TRANSFEM 4261 4280
 FT TRANSFEM 4281 4300
 FT TRANSFEM 4301 4320
 FT TRANSFEM 4321 4340
 FT TRANSFEM 4341 4360
 FT TRANSFEM 4361 4380
 FT TRANSFEM 4381 4400
 FT TRANSFEM 4401 4420
 FT TRANSFEM 4421 4440
 FT TRANSFEM 4441 4460
 FT TRANSFEM 4461 4480
 FT TRANSFEM 4481 4500
 FT TRANSFEM 4501 4520
 FT TRANSFEM 4521 4540
 FT TRANSFEM 4541 4560
 FT TRANSFEM 4561 4580
 FT TRANSFEM 4581 4600
 FT TRANSFEM 4601 4620
 FT TRANSFEM 4621 4640
 FT TRANSFEM 4641 4660
 FT TRANSFEM 4661 4680
 FT TRANSFEM 4681 4700
 FT TRANSFEM 4701 4720
 FT TRANSFEM 4721 4740
 FT TRANSFEM 4741 4760
 FT TRANSFEM 4761 4780
 FT TRANSFEM 4781 4800
 FT TRANSFEM 4801 4820
 FT TRANSFEM 4821 4840
 FT TRANSFEM 4841 4860
 FT TRANSFEM 4861 4880
 FT TRANSFEM 4881 4900
 FT TRANSFEM 4901 4920
 FT TRANSFEM 4921 4940
 FT TRANSFEM 4941 4960
 FT TRANSFEM 4961 4980
 FT TRANSFEM 4981 5000
 FT TRANSFEM 5001 5020
 FT TRANSFEM 5021 5040
 FT TRANSFEM 5041 5060
 FT TRANSFEM 5061 5080
 FT TRANSFEM 5081 5100
 FT TRANSFEM 5101 5120
 FT TRANSFEM 5121 5140
 FT TRANSFEM 5141 5160
 FT TRANSFEM 5161 5180
 FT TRANSFEM 5181 5200
 FT TRANSFEM 5201 5220
 FT TRANSFEM 5221 5240
 FT TRANSFEM 5241 5260
 FT TRANSFEM 5261 5280
 FT TRANSFEM 5281 5300
 FT TRANSFEM 5301 5320
 FT TRANSFEM 5321 5340
 FT TRANSFEM 5341 5360
 FT TRANSFEM 5361 5380
 FT TRANSFEM 5381 5400
 FT TRANSFEM 5401 5420
 FT TRANSFEM 5421 5440
 FT TRANSFEM 5441 5460
 FT TRANSFEM 5461 5480
 FT TRANSFEM 5481 5500
 FT TRANSFEM 5501 5520
 FT TRANSFEM 5521 5540
 FT TRANSFEM 5541 5560
 FT TRANSFEM 5561 5580
 FT TRANSFEM 5581 5600
 FT TRANSFEM 5601 5620
 FT TRANSFEM 5621 5640
 FT TRANSFEM 5641 5660
 FT TRANSFEM 5661 5680
 FT TRANSFEM 5681 5700
 FT TRANSFEM 5701 5720
 FT TRANSFEM 5721 5740
 FT TRANSFEM 5741 5760
 FT TRANSFEM 5761 5780
 FT TRANSFEM 5781 5800
 FT TRANSFEM 5801 5820
 FT TRANSFEM 5821 5840
 FT TRANSFEM 5841 5860
 FT TRANSFEM 5861 5880
 FT TRANSFEM 5881 5900
 FT TRANSFEM 5901 5920
 FT TRANSFEM 5921 5940
 FT TRANSFEM 5941 5960
 FT TRANSFEM 5961 5980
 FT TRANSFEM 5981 6000
 FT TRANSFEM 6001 6020
 FT TRANSFEM 6021 6040
 FT TRANSFEM 6041 6060
 FT TRANSFEM 6061 6080
 FT TRANSFEM 6081 6100
 FT TRANSFEM 6101 6120
 FT TRANSFEM 6121 6140
 FT TRANSFEM 6141 6160
 FT TRANSFEM 6161 6180
 FT TRANSFEM 6181 6200
 FT TRANSFEM 6201 6220
 FT TRANSFEM 6221 6240
 FT TRANSFEM 6241 6260
 FT TRANSFEM 6261 6280
 FT TRANSFEM 6281 6300
 FT TRANSFEM 6301 6320
 FT TRANSFEM 6321 6340
 FT TRANSFEM 6341 6360
 FT TRANSFEM 6361 6380
 FT TRANSFEM 6381 6400
 FT TRANSFEM 6401 6420
 FT TRANSFEM 6421 6440
 FT TRANSFEM 6441 6460
 FT TRANSFEM 6461 6480
 FT TRANSFEM 6481 6500
 FT TRANSFEM 6501 6520
 FT TRANSFEM 6521 6540
 FT TRANSFEM 6541 6560
 FT TRANSFEM 6561 6580
 FT TRANSFEM 6581 6600
 FT TRANSFEM 6601 6620
 FT TRANSFEM 6621 6640
 FT TRANSFEM 6641 6660
 FT TRANSFEM 6661 6680
 FT TRANSFEM 6681 6700
 FT TRANSFEM 6701 6720
 FT TRANSFEM 6721 6740
 FT TRANSFEM 6741 6760
 FT TRANSFEM 6761 6780
 FT TRANSFEM 6781 6800
 FT TRANSFEM 6801 6820
 FT TRANSFEM 6821 6840
 FT TRANSFEM 6841 6860
 FT TRANSFEM 6861 6880
 FT TRANSFEM 6881 6900
 FT TRANSFEM 6901 6920
 FT TRANSFEM 6921 6940
 FT TRANSFEM 6941 6960
 FT TRANSFEM 6961 6980
 FT TRANSFEM 6981 7000
 FT TRANSFEM 7001 7020
 FT TRANSFEM 7021 7040
 FT TRANSFEM 7041 7060
 FT TRANSFEM 7061 7080
 FT TRANSFEM 7081 7100
 FT TRANSFEM 7101 7120
 FT TRANSFEM 7121 7140
 FT TRANSFEM 7141 7160
 FT TRANSFEM 7161 7180
 FT TRANSFEM 7181 7200
 FT TRANSFEM 7201 7220
 FT TRANSFEM 7221 7240
 FT TRANSFEM 7241 7260
 FT TRANSFEM 7261 7280
 FT TRANSFEM 7281 7300
 FT TRANSFEM 7301 7320
 FT TRANSFEM 7321 7340
 FT TRANSFEM 7341 7360
 FT TRANSFEM 7361 7380
 FT TRANSFEM 7381 7400
 FT TRANSFEM 7401 7420
 FT TRANSFEM 7421 7440
 FT TRANSFEM 7441 7460
 FT TRANSFEM 7461 7480
 FT TRANSFEM 7481 7500
 FT TRANSFEM 7501 7520
 FT TRANSFEM 7521 7540
 FT TRANSFEM 7541 7560
 FT TRANSFEM 7561 7580
 FT TRANSFEM 7581 7600
 FT TRANSFEM 7601 7620
 FT TRANSFEM 7621 7640
 FT TRANSFEM 7641 7660
 FT TRANSFEM 7661 7680
 FT TRANSFEM 7681 7700
 FT TRANSFEM 7701 7720
 FT TRANSFEM 7721 7740
 FT TRANSFEM 7741 7760
 FT TRANSFEM 7761 7780
 FT TRANSFEM 7781 7800
 FT TRANSFEM 7801 7820
 FT TRANSFEM 7821 7840
 FT TRANSFEM 7841 7860
 FT TRANSFEM 7861 7880
 FT TRANSFEM 7881 7900
 FT TRANSFEM 7901 7920
 FT TRANSFEM 7921 7940
 FT TRANSFEM 7941 7960
 FT TRANSFEM 7961 7980
 FT TRANSFEM 7981 8000
 FT TRANSFEM 8001 8020
 FT TRANSFEM 8021 8040
 FT TRANSFEM 8041 8060
 FT TRANSFEM 8061 8080
 FT TRANSFEM 8081 8100
 FT TRANSFEM 8101 8120
 FT TRANSFEM 8121 8140
 FT TRANSFEM 8141 8160
 FT TRANSFEM 8161 8180
 FT TRANSFEM 8181 8200
 FT TRANSFEM 8201 8220
 FT TRANSFEM 8221 8240
 FT TRANSFEM 8241 8260
 FT TRANSFEM 8261 8280
 FT TRANSFEM 8281 8300
 FT TRANSFEM 8301 8320
 FT TRANSFEM 8321 8340
 FT TRANSFEM 8341 8360
 FT TRANSFEM 8361 8380
 FT TRANSFEM 8381 8400
 FT TRANSFEM 8401 8420
 FT TRANSFEM 8421 8440
 FT TRANSFEM 8441 8460
 FT TRANSFEM 8461 8480
 FT TRANSFEM 8481 8500
 FT TRANSFEM 8501 8520
 FT TRANSFEM 8521 8540
 FT TRANSFEM 8541 8560
 FT TRANSFEM 8561 8580
 FT TRANSFEM 8581 8600
 FT TRANSFEM 8601 8620
 FT TRANSFEM 8621 8640
 FT TRANSFEM 8641 8660
 FT TRANSFEM 8661 8680
 FT TRANSFEM 8681 8700
 FT TRANSFEM 8701 8720
 FT TRANSFEM 8721 8740
 FT TRANSFEM 8741 8760
 FT TRANSFEM 8761 8780
 FT TRANSFEM 8781 8800
 FT TRANSFEM 8801 8820
 FT TRANSFEM 8821 8840
 FT TRANSFEM 8841 8860
 FT TRANSFEM 8861 8880
 FT TRANSFEM 8881 8900
 FT TRANSFEM 8901 8920
 FT TRANSFEM 8921 8940
 FT TRANSFEM 8941 8960
 FT TRANSFEM 8961 8980
 FT TRANSFEM 8981 9000
 FT TRANSFEM 9001 9020
 FT TRANSFEM 9021 9040
 FT TRANSFEM 9041 9060
 FT TRANSFEM 9061 9080
 FT TRANSFEM 9081 9100
 FT TRANSFEM 9101 9120
 FT TRANSFEM 9121 9140
 FT TRANSFEM 9141 9160
 FT TRANSFEM 9161 9180
 FT TRANSFEM 9181 9200
 FT TRANSFEM 9201 9220
 FT TRANSFEM 9221 9240
 FT TRANSFEM 9241 9260
 FT TRANSFEM 9261 9280
 FT TRANSFEM 9281 9300
 FT TRANSFEM 9301 9320
 FT TRANSFEM 9321 9340
 FT TRANSFEM 9341 9360
 FT TRANSFEM 9361 9380
 FT TRANSFEM 9381 9400
 FT TRANSFEM 9401 9420
 FT TRANSFEM 9421 9440
 FT TRANSFEM 9441 9460
 FT TRANSFEM 9461 9480
 FT TRANSFEM 9481 9500
 FT TRANSFEM 9501 9520
 FT TRANSFEM 9521 9540
 FT TRANSFEM 9541 9560
 FT TRANSFEM 9561 9580
 FT TRANSFEM 9581 9600
 FT TRANSFEM 9601 9620
 FT TRANSFEM 9621 9640
 FT TRANSFEM 9641 9660
 FT TRANSFEM 9661 9680
 FT TRANSFEM 9681 9700
 FT TRANSFEM 9701 9720
 FT TRANSFEM 9721 9740
 FT TRANSFEM 9741 9760
 FT TRANSFEM 9761 9780
 FT TRANSFEM 9781 9800
 FT TRANSFEM 9801 9820
 FT TRANSFEM 9821 9840
 FT TRANSFEM 9841 9860
 FT TRANSFEM 9861 9880
 FT TRANSFEM 9881 9900
 FT TRANSFEM 9901 9920
 FT TRANSFEM 9921 9940
 FT TRANSFEM 9941 9960
 FT TRANSFEM 9961 9980
 FT TRANSFEM 9981 10000
 FT TRANSFEM 10001 10020
 FT TRANSFEM 10021 10040
 FT TRANSFEM 10041 10060
 FT TRANSFEM 10061 10080
 FT TRANSFEM 10081 10100
 FT TRANSFEM 10101 10120
 FT TRANSFEM 10121 10140
 FT TRANSFEM 10141 10160
 FT TRANSFEM 10161 10180
 FT TRANSFEM 10181 10200
 FT TRANSFEM 10201 10220
 FT TRANSFEM 10221 10240
 FT TRANSFEM 10241 10260
 FT TRANSFEM 10261 10280
 FT TRANSFEM 10281 10300
 FT TRANSFEM 10301 10320
 FT TRANSFEM 10321 10340
 FT TRANSFEM 10341 10360
 FT TRANSFEM 10361 10380
 FT TRANSFEM 10381 10400
 FT TRANSFEM 10401 10420
 FT TRANSFEM 10421 10440
 FT TRANSFEM 10441 10460
 FT TRANSFEM 10461 10480
 FT TRANSFEM 10481 10500
 FT TRANSFEM 10501 10520
 FT TRANSFEM 10521 10540
 FT TRANSFEM 10541 10560
 FT TRANSFEM 10561 10580
 FT TRANSFEM 10581 10600
 FT TRANSFEM 10601 10620
 FT TRANSFEM 10621 10640
 FT TRANSFEM 10641 10660
 FT TRANSFEM 10661 10680
 FT TRANSFEM 10681 10700
 FT TRANSFEM 10701 10720
 FT TRANSFEM 10721 10740
 FT TRANSFEM 10741 10760
 FT TRANSFEM 10761 10780
 FT TRANSFEM 10781 10800
 FT TRANSFEM 10801 10820
 FT TRANSFEM 10821 10840
 FT TRANSFEM 10841 10860
 FT TRANSFEM 10861 10880
 FT TRANSFEM 10881 10900
 FT TRANSFEM 10901 10920
 FT TRANSFEM 10921 10940
 FT TRANSFEM 10941 10960
 FT TRANSFEM 10961 10980
 FT TRANSFEM 10981 11000
 FT TRANSFEM 11001 11020
 FT TRANSFEM 11021 11040
 FT TRANSFEM 11041 11060
 FT TRANSFEM 11061 11080
 FT TRANSFEM 11081 11100
 FT TRANSFEM 11101 11120
 FT TRANSFEM 11121 11140
 FT TRANSFEM 11141 11160
 FT TRANSFEM 11161 11180
 FT TRANSFEM 11181 11200
 FT TRANSFEM 11201 11220
 FT TRANSFEM 11221 11240
 FT TRANSFEM 11241 11260
 FT TRANSFEM 11261 11280
 FT TRANSFEM 11281 11300
 FT TRANSFEM 11301 11320
 FT TRANSFEM 11321 11340
 FT TRANSFEM 11341 11360
 FT TRANSFEM 11361 11380
 FT TRANSFEM 11381 11400
 FT TRANSFEM 11401 11420
 FT TRANSFEM 11421 11440
 FT TRANSFEM 11441 11460
 FT TRANSFEM 11461 11480
 FT TRANSFEM 11481 11500
 FT TRANSFEM 11501 11520
 FT TRANSFEM 11521 11540
 FT TRANSFEM 11541 11560
 FT TRANSFEM 11561 11580
 FT TRANSFEM 11581 11600
 FT TRANSFEM 11601 11620
 FT TRANSFEM 11621 11640
 FT TRANSFEM 11641 11660
 FT TRANSFEM 11661 11680
 FT TRANSFEM 11681 11700
 FT TRANSFEM 11701 11720
 FT TRANSFEM 11721 11740
 FT TRANSFEM 11741 11760
 FT TRANSFEM 11761 11780
 FT TRANSFEM 11781 11800
 FT TRANSFEM 11801 11820
 FT TRANSFEM 11821 11840
 FT TRANSFEM 11841 11860
 FT TRANSFEM 11861 11880
 FT TRANSFEM 11881 11900
 FT TRANSFEM 11901 11920
 FT TRANSFEM 11921 11940
 FT TRANSFEM 11941 11960
 FT TRANSFEM 11961 11980
 FT TRANSFEM 11981 12000
 FT TRANSFEM 12001 12020
 FT TRANSFEM 12021 12040
 FT TRANSFEM 12041 12060
 FT TRANSFEM 12061 12080
 FT TRANSFEM 12081 12100
 FT TRANSFEM 12101 12120
 FT TRANSFEM 12121 12140
 FT TRANSFEM 12141 12160
 FT TRANSFEM 12161 12180
 FT TRANSFEM 12181 12200
 FT TRANSFEM 12201 12220
 FT TRANSFEM 12221 12240
 FT TRANSFEM 12241 12260
 FT TRANSFEM 12261 12280
 FT TRANSFEM 12281 12300
 FT TRANSFEM 12301 12320
 FT TRANSFEM 12321 12340
 FT TRANSFEM 12341 12360
 FT TRANSFEM 12361 12380
 FT TRANSFEM 12381 12400
 FT TRANSFEM 12401 12420
 FT TRANSFEM 12421 12440
 FT TRANSFEM 12441 12460
 FT TRANSFEM 12461 12480
 FT TRANSFEM 12481 12500
 FT TRANSFEM 12501 12520
 FT TRANSFEM 12521 12540
 FT TRANSFEM 12541 12560
 FT TRANSFEM 12561 12580

236 DPOVGVGVDQVQLNKYDSWISFLSSVRYVMWAFNVRACQSYGCGVCISGPLG 289

Db

RESULT 13

Q8CEB9

ID Q8CEB9 PRELIMINARY; PRT; 554 AA.

AC Q8CEB9;

DT 01-MAR-2003 (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched

DE library, clone:4732404L04 product:similar to D642111.

GN Name=Haas3;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

RN [1]

SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=Skin;

RP MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RX Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RT Meth. Enzymol. 303:19-44(1999).

RN [2]

SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=Skin;

RP MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RX RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RT Nature 409:685-690(2001).

RN [3]

SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=Skin;

RP The FANTOM Consortium;

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RT Nature 420:563-573(2002).

RN [4]

SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=Skin;

RP MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RX Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RT Genome Res. 10:1617-1630(2000).

RN [5]

SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=Skin;

RP MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,

RA Sumi N., Ishii Y., Nakamura S., Haghama M., Nishine T., Harada A.,

RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format

RT sequencing pipeline with 384 multipipillary sequeencer.";

RT Genome Res. 10:1757-1771(2000).

RN [6]

SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=Skin;

RP Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,

RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,

RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,

RA Kato H., Kawai J., Koijima Y., Kondo S., Konno H., Kouada M., Koya S.,

RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,

RA Nishi K., Nomura K., Numasaki R., Ohno M., Ohsato N., Okazaki Y.,

RA Saito R., Saichoh H., Sakai C., Sakai K., Sakazume N., Sano H.,

RN [7]

```
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK028582; BAC26017.1; -.
DR MGD; MGI:109599; Haas3.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR InterPro; IPR001173; Glyco_trans_2.
SQ SEQUENCE 554 AA; 63348 MW; BFF239C901251DB CRC64;

Query Match 39.8%; Score 791; DB 2; Length 554;
Best Local Similarity 55.4%; Pred. No. 2.4e-53;
Matches 163; Conservative 36; Mismatches 77; Indels 18; Gaps 5;

Qy 21 RRVLTAFALLIILGLMTWAYAAGVPLASDRYGLLAFLYGAFLSAHLVAQSLFAYLEHRR 80
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
10 RVVGTSLFALVWLGILAAVVTGYQFIHTEKHLSFGLYGAILGLHLIIQSLFAFLHRR 69

Qy 81 VAAARGPLD---AATARSVALTISAYQEDPAYLRQCILASARALLYPRALRLVLMVVDGN 137
Db ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
70 MRRAGR-PLKLHCSQRPRRSVALCIAAYQEDPEYLRKCLRSQAQRIAPP--NLKVVVMVVDGN 126

Qy 138 RAEDLYVMDRFVVF-ADEDPATVYWDGNHQPWEPAAAGAVGAGAYREVEAEDPGRLAV 196
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
127 RQEDTYMLDIFHEVLGTEQAGFFWRSNFHEAGEGETEASLQEGMER-----V 175

Qy 197 EALVTRRCVCVAQRWGGKREVMYTAFAKALGSDSYDVYQVCDSDTRLDPMALLELVRVLE 256
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
176 RAVVWASTFSCIMQKMGKREVMYTAFAKALGNSVDYIQVCDSDTVLDPACTIEMLRVLEE 235

Qy 257 DPRVAGVGDVRILNPLDSWVSLSSRLRYVAFNVERACOSYFHCVCISGSLG 310
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
236 DPQVGGVGDVQILNKYDWSISFLSSRYVWAFNVERACOSYFGVCQICISGPLG 289

RESULT 15
Q8CH92 ID Q8CH92 PRELIMINARY; PRT; 554 AA.
AC Q8CH92;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hyaluronan synthase 3.
GN Name=HAS3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14724275; DOI=10.1074/jbc.M313178200;
RA Itano N., Sawai T., Ateumi F., Miyaishi O., Taniguchi S., Kannagi R.,
RA Hamaguchi M., Kimata K.;
RT "Selective expression and functional characteristics of three
RT Mammalian hyaluronan synthases in oncogenic malignant
RT transformation.";
RL J. Biol. Chem. 279:18679-18687 (2004).
DR EMBL; AB097569; BAC43731.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
SQ SEQUENCE 554 AA; 63349 MW; 475D5BD62717CE51 CRC64;

Query Match 39.8%; Score 791; DB 2; Length 554;
Best Local Similarity 55.4%; Pred. No. 2.4e-53;
Matches 163; Conservative 36; Mismatches 77; Indels 18; Gaps 5;

Qy 21 RRVLTAFALLIILGLMTWAYAAGVPLASDRYGLLAFLYGAFLSAHLVAQSLFAYLEHRR 80
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
10 RVVGTSLFALVWLGILAAVVTGYQFIHTEKHLSFGLYGAILGLHLIIQSLFAFLHRR 69

Qy 81 VAAARGPLD---AATARSVALTISAYQEDPAYLRQCILASARALLYPRALRLVLMVVDGN 137
Db ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
70 MRRAGR-PLKLHCSQRPRRSVALCIAAYQEDPEYLRKCLRSQAQRIAPP--NLKVVVMVVDGN 126
```

```
Qy 138 RAEDLYVMDRFVVF-ADEDPATVYWDGNHQPWEPAAAGAVGAGAYREVEAEDPGRLAV 196
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
127 RQEDTYMLDIFHEVLGTEQAGFFWRSNFHEAGEGETEASLQEGMER-----V 175

Qy 197 EALVTRRCVCVAQRWGGKREVMYTAFAKALGSDSYDVYQVCDSDTRLDPMALLELVRVLE 256
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
176 RAVVWASTFSCIMQKMGKREVMYTAFAKALGNSVDYIQVCDSDTVLDPACTIEMLRVLEE 235

Qy 257 DPRVAGVGDVRILNPLDSWVSLSSRLRYVAFNVERACOSYFHCVCISGSLG 310
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
236 DPQVGGVGDVQILNKYDWSISFLSSRYVWAFNVERACOSYFGVCQICISGPLG 289

RESULT 15
Q9DG40 ID Q9DG40 PRELIMINARY; PRT; 554 AA.
AC Q9DG40;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hyaluronan synthase 3.
GN Name=has3;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14729574;
RA Bakkens J., Kramer C., Pothof J., Quaedvlieg N.E., Spaink H.P.,
RA Hammerschmidt M.;
RT "Has2 is required upstream of Rac1 to govern dorsal migration of
RT lateral cells during zebrafish gastrulation.";
RL Development 131:525-537 (2004).
DR EMBL; AF190743; AAG17181.1; -.
DR ZFIN; ZDB-GENE-021118-1; has3.
DR InterPro; IPR001173; Glyco_trans_2.
SQ SEQUENCE 554 AA; 63077 MW; 26273CB459CE3E97 CRC64;

Query Match 39.8%; Score 790.5; DB 2; Length 554;
Best Local Similarity 55.7%; Pred. No. 2.6e-53;
Matches 165; Conservative 40; Mismatches 74; Indels 17; Gaps 8;

Qy 18 GLARRV-LTIAFALLIILGLMTWAYAAGVPLASDRYGLLAFLYGAFLSAHLVAQSLFAYL 76
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
6 GTAVRPIITLFAAVLVFAILLAYVTGYQFIHTEKHLSFGLYGAILGLHLIIQSLFAYL 65

Qy 77 EHRVAAARGP-LDAATARSVALTISAYQEDPAYLRQCILASARALLYPRALRLVLMVVD 135
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
66 EHRQM----RGRSRPQHLRTRTVALCIAAYQEDPYLRKCLRSSR-ISFP--GLKVLIVVD 118

Qy 136 GNAEADLYVMDRFVVFAD-EDPATVYWDGNHQPWEPAAAGAVGAGAYREVEAEDPGLR 194
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
119 GNRQEDAYMDDIFQEVWVGVEQTCVWKNYHNSGDDGGGGKGS-----VHAEEARV 173

Qy 195 AVBALVTRRCVCVCAQRWGGKREVMYTAFAKALGSDSYDVYQVCDSDTRLDPMALLELVRV 254
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
174 A--RVVRSRCRYSCIMQEWGGKREVMYTAFAKALGDTVDYMQVCDSDTVLDPACTIEMLRKIL 231

Qy 255 DEDPRVAGVGDVRILNPLDSWVSLSSRLRYVAFNVERACOSYFHCVCISGSLG 310
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
232 EEDPDVGGVGDVQILNKYDWSISFLSSRYVWAFNVERACOSYFGVCQICISGPLG 287

Search completed: March 11, 2005, 14:21:56
Job time : 100.273 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 08:01:42 ; Search time 753.899 Seconds
(without alignments)
8450.560 Million cell updates/sec

Title: US-10-672-399-3

Perfect score: 1071

Sequence: 1 atgagacagcaggacgcgc.....ccggcaccatcggtgatga 1071

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09D_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1071	100.0	1071	18	US-10-672-399-3
2	953.4	89.0	1083	18	US-10-672-399-5
3	927.4	86.6	1065	18	US-10-672-399-7
4	927.4	86.6	1737	18	US-10-672-399-1
5	927.4	86.6	2116	13	US-10-042-523-1
6	678.2	63.3	1752	10	US-09-902-939-1
7	422	39.4	662	18	US-10-363-345A-13975
8	422	39.4	662	18	US-10-363-345A-13976
9	393.8	36.8	662	18	US-10-363-345A-13973
10	393.8	36.8	662	18	US-10-363-345A-13974
11	338.6	31.6	490	10	US-09-918-995-27210

12	272.8	25.5	1662	17	US-10-295-027-369	Sequence 369, App
13	272.8	25.5	1662	17	US-10-188-832-136	Sequence 136, App
14	248.4	23.2	1767	17	US-10-309-560-9	Sequence 9, Appli
15	236.2	22.1	1665	10	US-09-902-939-3	Sequence 3, Appli
16	176.2	16.5	1659	10	US-09-902-939-2	Sequence 2, Appli
17	176.2	16.5	4194	15	US-10-262-526-3	Sequence 3, Appli
18	148.8	13.9	3003	10	US-09-918-624B-64	Sequence 64, Appli
19	148.8	13.9	3003	15	US-10-262-526-1	Sequence 1, Appli
20	148.8	13.9	3387	15	US-10-084-817-184	Sequence 184, App
21	148.8	13.9	3387	16	US-10-252-157-84	Sequence 84, Appli
22	148.8	13.9	4018	10	US-09-814-353-19942	Sequence 15942, A
23	100.2	9.4	793	9	US-09-910-943-552	Sequence 552, App
24	81.4	7.6	570	9	US-09-880-107-349	Sequence 349, Ap
25	72	6.7	15559	18	US-10-646-664-1	Sequence 1, Appli
26	71	6.6	2182	18	US-10-437-963-62269	Sequence 62269, A
27	70	6.5	43058	9	US-09-954-456-292	Sequence 292, App
28	70	6.5	43058	9	US-09-954-456-529	Sequence 529, App
29	70	6.5	43058	9	US-09-880-107-3950	Sequence 3950, Ap
30	69.4	6.5	5784	14	US-10-152-886-64	Sequence 64, Appli
31	68.6	6.4	1755	18	US-10-437-963-78301	Sequence 78301, A
32	68.4	6.4	2214	15	US-10-156-761-2510	Sequence 2510, Ap
33	68.4	6.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
34	67.6	6.3	1380	18	US-10-425-115-170699	Sequence 170699
35	66.6	6.2	3957	16	US-10-200-562-193	Sequence 193, App
36	66.6	6.2	3957	16	US-10-237-551-193	Sequence 193, App
37	66.6	6.2	154746	10	US-09-827-688-8	Sequence 8, Appli
38	66.6	6.2	154746	10	US-09-827-688-8	Sequence 8, Appli
39	66.2	6.2	1614	9	US-09-976-740-45	Sequence 45, Appli
40	66.2	6.2	1614	13	US-10-023-529-45	Sequence 45, Appli
41	66.2	6.2	1614	13	US-10-023-523-45	Sequence 45, Appli
42	66.2	6.2	1614	17	US-10-616-187-45	Sequence 45, Appli
43	66.2	6.2	1614	17	US-10-671-242-45	Sequence 45, Appli
44	66.2	6.2	12425	9	US-09-976-740-50	Sequence 50, Appli
45	66.2	6.2	12425	13	US-10-023-529-50	Sequence 50, Appli

ALIGNMENTS

RESULT 1

US-10-672-399-3
; Sequence 3, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-672-399-3

Query Match	100.0%	Score 1071	DB 18	Length 1071
Best Local Similarity	100.0%	Pred. No. 1.1e-256		
Matches 1071	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCGGCTTGCC 60		
Db	1	ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCGGCTTGCC 60		
Qy	61	CGGAGGGTGTGACCATCGCCTTCGCTGCTATCTCTGGCCCTCATGACCTGGGCTTAC 120		
Db	61	CGGAGGGTGTGACCATCGCCTTCGCTGCTATCTCTGGCCCTCATGACCTGGGCTTAC 120		
Qy	121	GCCCGCGGGTGCCTGGCTTCCGATCGCTACGGCTTCTGGGCTTCTGCGCTCTACCGG 180		
Db	121	GCCCGCGGGTGCCTGGCTTCCGATCGCTACGGCTTCTGGGCTTCTGCGCTCTACCGG 180		

Db 121 GCCCGCGGGTGCCGCTCCGATCGCTACGGCCTCCTGGCCTTCAGCGG 180
Qy 181 GCCTTCCTTTAGCGCACTGGTGGCGAGAGCCTTTCGGGTACTTGGAGCAACGGCGG 240
Db 181 GCCTTCCTTTAGCGCACTGGTGGCGAGAGCCTTTCGGGTACTTGGAGCAACGGCGG 240
Qy 241 GTGG 300
Db 241 GTGG 300
Qy 301 ATCTCCGCTTACCAAGAGAGACCCCGCTACCTGGCCAGTGCCTGGCGCTCCGCGCGCC 360
Db 301 ATCTCCGCTTACCAAGAGAGACCCCGCTACCTGGCCAGTGCCTGGCGCTCCGCGCGCC 360
Qy 361 CTGCTGTATCCCGCGCGCGCTGGCGCTCTCATGTGTGGTATGGCAACCGCGCCGAG 420
Db 361 CTGCTGTATCCCGCGCGCGCTGGCGCTCTCATGTGTGGTATGGCAACCGCGCCGAG 420
Qy 421 GACCTCTACATGGTGCAGATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCAGTAC 480
Db 421 GACCTCTACATGGTGCAGATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCAGTAC 480
Qy 481 GTGTGGGACGCAACTACCAAGCCCTGGGAAACCGCGCGCGGGCGCGGTGGGGGCC 540
Db 481 GTGTGGGACGCAACTACCAAGCCCTGGGAAACCGCGCGCGGGCGCGGTGGGGGCC 540
Qy 541 GGAGCCTATCCGGAGGTGGAGGGAGGATCCTTGGCGGGCTTGGCAGTGGAGGCGCTGGTG 600
Db 541 GGAGCCTATCCGGAGGTGGAGGGAGGATCCTTGGCGGGCTTGGCAGTGGAGGCGCTGGTG 600
Qy 601 AGGACTCCAGGTGGGTGGCGCGAGCGCTGGGGCGGCGCAACCGCGAGGTCTATGTAC 660
Db 601 AGGACTCCAGGTGGGTGGCGCGAGCGCTGGGGCGGCGCAACCGCGAGGTCTATGTAC 660
Qy 661 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGACTACGTGAGGTCTGTGACTCGGACACA 720
Db 661 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGACTACGTGAGGTCTGTGACTCGGACACA 720
Qy 721 AGGTTGGAACCCATGGCACTGCTGGAGCTGTGGCGGGTACTGGAAGGACCCCGGGTA 780
Db 721 AGGTTGGAACCCATGGCACTGCTGGAGCTGTGGCGGGTACTGGAAGGACCCCGGGTA 780
Qy 781 GGGGCTGTGTGGGAGCGTGGATCCTTAACCTCTGGACTCTGGTTCAGGCTTCTTA 840
Db 781 GGGGCTGTGTGGGAGCGTGGATCCTTAACCTCTGGACTCTGGTTCAGGCTTCTTA 840
Qy 841 AGCAGCCTGCATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTTCAGAGCTACTTCCAC 900
Db 841 AGCAGCCTGCATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTTCAGAGCTACTTCCAC 900
Qy 901 TGTGTATCTGCATCAGCGGTTCTTAGGTACACTTCAAGTCCCGCTGCTTCAAGAG 960
Db 901 TGTGTATCTGCATCAGCGGTTCTTAGGTACACTTCAAGTCCCGCTGCTTCAAGAG 960
Qy 961 CGCCTCGTCTTCTGGGGTGGTGGAGCAGGAGACAGCGTGGTCCAGTGGTACTTCC 1020
Db 961 CGCCTCGTCTTCTGGGGTGGTGGAGCAGGAGACAGCGTGGTCCAGTGGTACTTCC 1020
Qy 1021 GTGAGTGGCTGACAAAGCGCTCTGGGTGGACCCGGCACCATGGCTGGATGA 1071
Db 1021 GTGAGTGGCTGACAAAGCGCTCTGGGTGGACCCGGCACCATGGCTGGATGA 1071

RESULT 2

US-10-672-399-5
; Sequence 5, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A994635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25

; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-5

Query Match 89.0%; Score 953.4; DB 18; Length 1083;
Beet Local Similarity 94.5%; Pred. No. 1.9e-227;
Matches 1023; Conservative 0; Mismatches 1; Indels 59; Gaps 1;

Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTTCAGCCCGCCCGCTGTCTCGGCTTGCGC 60
Db 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTTCAGCCCGCCCGCTGTCTCGGCTTGCGC 60
Qy 61 CGGAGGGTGTGACCATCGGCTTCGCCCTTCGCTATCTTCCTGGGCTTCATGACCTGGGCTTAC 120
Db 61 CGGAGGGTGTGACCATCGGCTTCGCCCTTCGCTATCTTCCTGGGCTTCATGACCTGGGCTTAC 120
Qy 121 GCCCGCGGGGTGCCGCTGGGCTTCGATCGCTTACGCGCTCTCTGGGCTTCCTACGGG 180
Db 121 GCCCGCGGGGTGCCGCTGGGCTTCGATCGCTTACGCGCTCTCTGGGCTTCCTACGGG 180
Qy 181 GCCTTCCTTTACGCGCACTTGGTGGCGCAGAGCCTTTCGGGTACTTGGAGCAACGGCGG 240
Db 181 GCCTTCCTTTACGCGCACTTGGTGGCGCAGAGCCTTTCGGGTACTTGGAGCAACGGCGG 240
Qy 241 GTGGCGCGCGCGCGCGGGGGCGCTGGATGCGACCAACCGCGCGAGTGTGGCGCTGACC 300
Db 241 GTGGCGCGCGCGCGCGGGGGCGCTGGATGCGACCAACCGCGCGAGTGTGGCGCTGACC 300
Qy 301 ATCTCCCGCTTACCAAGAGAGACCCCGCTACCTGGCGCAGTGCCTGGCGCTCCGCGCGCC 360
Db 301 ATCTCCCGCTTACCAAGAGAGACCCCGCTACCTGGCGCAGTGCCTGGCGCTCCGCGCGCC 360
Qy 361 CTGCTGTATCCCGCGCGCGCTGGCGCTCTCATGTGTGGTATGGCAACCGCGCCGAG 420
Db 361 CTGCTGTATCCCGCGCGCGCTGGCGCTCTCATGTGTGGTATGGCAACCGCGCCGAG 420
Qy 421 GACCTCTACATGGTGCAGATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCAGTAC 480
Db 421 GACCTCTACATGGTGCAGATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCAGTAC 480
Qy 481 GTGTGGGACGCAACTACCAAGCCCTGGGAAACCGCGCGCGGGCGCGGTGGGGGCC 540
Db 481 GTGTGGGACGCAACTACCAAGCCCTGGGAAACCGCGCGCGGGCGCGGTGGGGGCC 540
Qy 541 GGAGCCTATCCGGAGGTGGAGGGAGGATCCTTGGCGGGCTTGGCAGTGGAGGCGCTGGTG 600
Db 541 GGAGCCTATCCGGAGGTGGAGGGAGGATCCTTGGCGGGCTTGGCAGTGGAGGCGCTGGTG 600
Qy 601 AGGACTCCAGGTGGGTGGCGCGAGCGCTGGGGCGGCGCAACCGCGAGGTCTATGTAC 660
Db 601 AGGACTCCAGGTGGGTGGCGCGAGCGCTGGGGCGGCGCAACCGCGAGGTCTATGTAC 660
Qy 661 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGACTACGTGAGGTCTGTGACTCGGACACA 720
Db 661 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGACTACGTGAGGTCTGTGACTCGGACACA 720
Qy 721 AGGTTGGAACCCATGGCACTGCTGGAGCTGTGGCGGGTACTGGAAGGACCCCGGGTA 780
Db 721 AGGTTGGAACCCATGGCACTGCTGGAGCTGTGGCGGGTACTGGAAGGACCCCGGGTA 780
Qy 781 GGGGCTGTGTGGGAGCGTGGATCCTTAACCTCTGGACTCTGGTTCAGGCTTCTTA 840
Db 781 GGGGCTGTGTGGGAGCGTGGATCCTTAACCTCTGGACTCTGGTTCAGGCTTCTTA 840
Qy 841 AGCAGCCTGCATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTTCAGAGCTACTTCCAC 900
Db 841 AGCAGCCTGCATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTTCAGAGCTACTTCCAC 900

```
QY 901 TGTGTATCTGTGATCAGCGTTCTCT----- 926
Db 901 TGTGTATCTGTGATCAGCGTCTCTAGAACTCTGCCAGGCCCCAGGAGCAGCGATG 960
QY 927 -----AGGTACACCTCCAGTCCCTGCTACTCAGAC 961
Db 961 ATGCCCTCATCTCTCCGCCCGGTGAGGTACACCTCCAGGTCCCGTCTGCTACTCAGAGAC 1020
QY 962 GCCCTCGTCTCTCTCCGCGTGTGCTGAGCCAGCAGACAGCTGCTCAAGTCTGCTACTTCG 1021
Db 1021 GCCCTCGTCTCTCTCCGCGTGTGCTGAGCCAGCAGACAGCTGCTCAAGTCTGCTACTTCG 1080
QY 1022 TGA 1024
Db 1081 TGA 1083

RESULT 3
US-10-672-399-7
; Sequence 7, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-7

Query Match 86.6%; Score 927.4; DB 18; Length 1065;
Best Local Similarity 99.9%; Pred. No. 5.5e-221;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTGTGAGCCGCGCGCTGTCTCGGCTTGCC 60
Db 1 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTGTGAGCCGCGCGCTGTCTCGGCTTGCC 60
QY 61 CGGAGGGTGTGACCATCGCCTTCCGCTGTCTGCTGAGCCCTCATGACCTGGGCTTAC 120
Db 61 CGGAGGGTGTGACCATCGCCTTCCGCTGTCTGCTGAGCCCTCATGACCTGGGCTTAC 120
QY 121 GCCCGCGGGGTGCGCGTCCGATCGCTACGCGCCTCTGGCCTTCCGCTCTACCGG 180
Db 121 GCCCGCGGGGTGCGCGTCCGATCGCTACGCGCCTCTGGCCTTCCGCTCTACCGG 180
QY 181 GCCTTCTTTTACGCGACCTGTGTGCGCAGAGCCTCTTTCGCTACGAGACCCCGCCAG 240
Db 181 GCCTTCTTTTACGCGACCTGTGTGCGCAGAGCCTCTTTCGCTACGAGACCCCGCCAG 240
QY 241 GTGGCGGCGGGCGCGCGGCGCGCTGATGCGACGACGCGCGCAGTGTGGCGCTGACC 300
Db 241 GTGGCGGCGGGCGCGCGGCGCGCTGATGCGACGACGCGCGCAGTGTGGCGCTGACC 300
QY 301 ATCTCCGCTTACACGAGAGACCCCGCTACCTGCGCAGTGTGCTGCGCGCGCGCC 360
Db 301 ATCTCCGCTTACACGAGAGACCCCGCTACCTGCGCAGTGTGCTGCGCGCGCGCC 360
QY 361 CTGCTGTATACCGCGCGCGCGTGTGCGCTCTCTATGTGTGTGATGCGCAACCGCGCCAG 420
Db 361 CTGCTGTATACCGCGCGCGCGTGTGCGCTCTCTATGTGTGTGATGCGCAACCGCGCCAG 420
QY 421 GACCTCTACATGGTTCGATGTTCCGCGAGGTCTTTCGCTGACGAGACCCCGCCAGTAC 480
Db 421 GACCTCTACATGGTTCGATGTTCCGCGAGGTCTTTCGCTGACGAGACCCCGCCAGTAC 480
```

```
QY 481 GTGTGGGACGGCAACTACACAGCCCTTGGAAACCCGCGCGCGCGCGCGTGGCGGCC 540
Db 481 GTGTGGGACGGCAACTACACAGCCCTTGGAAACCCGCGCGCGCGCGCGTGGCGGCC 540
QY 541 GGAGCCTATCCGAGGTGAGCGCGAGATCTCTGGGCGGCTGGCAGTGTGAGGCGCTGGTG 600
Db 541 GGAGCCTATCCGAGGTGAGCGCGAGATCTCTGGGCGGCTGGCAGTGTGAGGCGCTGGTG 600
QY 601 AGGACTCCAGGTGCTGTGCGCGCAGCGCTGGGCGGCAAGCGCGAGGTCAATGTAC 660
Db 601 AGGACTCCAGGTGCTGTGCGCGCAGCGCTGGGCGGCAAGCGCGAGGTCAATGTAC 660
QY 661 ACAGCCTTCAAGCGGCTCGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACA 720
Db 661 ACAGCCTTCAAGCGGCTCGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACA 720
QY 721 AGTTGGACCCCATGGGACTCTGTGGAGCTCTGGGGTACTGCGACGAGACCCCGGGTA 780
Db 721 AGTTGGACCCCATGGGACTCTGTGGAGCTCTGGGGTACTGCGACGAGACCCCGGGTA 780
QY 781 GGGGCTGTGTGGGAGCTCGGATCCTTAAACCTCTGGACTCTCTGGTCAAGCTTCTTA 840
Db 781 GGGGCTGTGTGGGAGCTCGGATCCTTAAACCTCTGGACTCTCTGGTCAAGCTTCTTA 840
QY 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTACAGCTTACTTCCAC 900
Db 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTACAGCTTACTTCCAC 900
QY 901 TGTGTATCTGCATCAGCGGTTCTCTAGG 929
Db 901 TGTGTATCTGCATCAGCGGTTCTCTAGG 929
```

RESULT 4

US-10-672-399-1

; Sequence 1, Application US/10672399

; Publication No. US20050003368A1

; GENERAL INFORMATION:

; APPLICANT: University of Alberta

; TITLE OF INVENTION: Cancer Monitoring and Therapeutics

; FILE REFERENCE: A894635US

; CURRENT APPLICATION NUMBER: US/10/672,399

; CURRENT FILING DATE: 2003-09-25

; PRIOR APPLICATION NUMBER: US 60/472,401

; PRIOR FILING DATE: 2003-05-22

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 1737

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-672-399-1

Query Match 86.6%; Score 927.4; DB 18; Length 1737;
Best Local Similarity 99.9%; Pred. No. 5.4e-221;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTGTGAGCCGCGCGCTGTCTCGGCTTGCC 60
Db 1 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTGTGAGCCGCGCGCTGTCTCGGCTTGCC 60
QY 61 CGGAGGGTGTGACCATCGCCTTCCGCTGTCTGCTGAGCCCTCATGACCTGGGCTTAC 120
Db 61 CGGAGGGTGTGACCATCGCCTTCCGCTGTCTGCTGAGCCCTCATGACCTGGGCTTAC 120
QY 121 GCCCGCGGGGTGCGCGTCCGATCGCTACGCGCCTCTGGCCTTCCGCTCTACCGG 180
Db 121 GCCCGCGGGGTGCGCGTCCGATCGCTACGCGCCTCTGGCCTTCCGCTCTACCGG 180
QY 181 GCCTTCTTTTACGCGACCTGTGTGCGCAGAGCCTTTCGCTGACGAGACCCCGCCAG 240
Db 181 GCCTTCTTTTACGCGACCTGTGTGCGCAGAGCCTTTCGCTGACGAGACCCCGCCAG 240
```


QY	241	GTGGCGGCGCGCGCGGGGCCGTGGATGCAGCCACCGCGCAGTGTGGCCTEACC	300
DB	241	GTGGCGGCGCGCGCGGGGCCGTGGATGCAGCCACCGCGCAGTGTGGCGCTGACC	300
QY	301	ATCTCCGCTTACCAGGAGACCCCGCTACCTGGCCAGTGCCTGGCGTCCGCCCGGCC	360
DB	301	ATCTCCGCTTACCAGGAGACCCCGCTACCTGGCCAGTGCCTGGCGTCCGCCCGGCC	360
QY	361	CTGCTGTACCCGCGCGCGGCTGGCGCTCTCATGTGTGTGATGGCAACCGCGCCGAG	420
DB	361	CTGCTGTACCCGCGCGCGGCTGGCGCTCTCATGTGTGTGATGGCAACCGCGCCGAG	420
QY	421	GACCTCTACATGGTGCACATGTTCCGCGAGGCTTTCGTGACGAGGACCCCGCCACGTAC	480
DB	421	GACCTCTACATGGTGCACATGTTCCGCGAGGCTTTCGTGACGAGGACCCCGCCACGTAC	480
QY	481	GTGTGGGACGGCAATTACCACAGCCCTGGGAAACCGCGCGCGCGCGGTGGCGGCC	540
DB	481	GTGTGGGACGGCAATTACCACAGCCCTGGGAAACCGCGCGCGCGCGGTGGCGGCC	540
QY	541	GGAGCCTATCGGAGGTGGAGCGGAGGATCCTTGGCGGGCTGGCAGTGGAGGCGCTGGTG	600
DB	541	GGAGCCTATCGGAGGTGGAGCGGAGGATCCTTGGCGGGCTGGCAGTGGAGGCGCTGGTG	600
QY	601	AGGACTCGCAGGTGTGTGGCGCAGCGCTGGGGCGGCAAGCCGCGAGGTCATGTAC	660
DB	601	AGGACTCGCAGGTGTGTGGCGCAGCGCTGGGGCGGCAAGCCGCGAGGTCATGTAC	660
QY	661	ACAGCCTTCAAGCGCTCGGAGATTCGGTGACGTACGTGCAGGTCGTGTACTCGGACACA	720
DB	661	ACAGCCTTCAAGCGCTCGGAGATTCGGTGACGTACGTGCAGGTCGTGTACTCGGACACA	720
QY	721	AGGTTGGACCCCATGGCACTGCTGAGCTCGTGGGGGTACTGGACGAGGACCCCGGGTA	780
DB	721	AGGTTGGACCCCATGGCACTGCTGAGCTCGTGGGGGTACTGGACGAGGACCCCGGGTA	780
QY	781	GGGGCTGTTGTTGGGACGTCGGGATCCTTAAACCTCTGGACTCCTGGGTACGCTTCCTA	840
DB	781	GGGGCTGTTGTTGGGACGTCGGGATCCTTAAACCTCTGGACTCCTGGGTACGCTTCCTA	840
QY	841	AGCAGCTCCGATATCTGGGTACGCTTCAATGTGGAGCGGGCTGTGTAGAGCTACTTCCAC	900
DB	841	AGCAGCTCCGATATCTGGGTACGCTTCAATGTGGAGCGGGCTGTGTAGAGCTACTTCCAC	900
QY	901	TGTGTATCTTCGATCAGCGGTTCTTAGG	929
DB	901	TGTGTATCTTCGATCAGCGGTTCTTAGG	929

RESULT 5

```

US-10-042-523-1
; Sequence 1, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523

```


QY 661 ACAGCCCTTCAGGCGCTCGAGATTCGGTGGACTACGTGCGAGTCTGTGACTCGGACACA 720
Db 696 ACAGCCCTTCAGGCGCTCGAGATTCGGTGGACTACGTGCGAGTCTGTGACTCGGACACA 755
QY 721 AGGTTGGACCCCATGGCACTGCTGGAGCTCGTGGGGTACTGGACGAGGACCCCGGGTA 780
Db 756 AGGTTGGACCCCATGGCACTGCTGGAGCTCGTGGGGTACTGGACGAGGACCCCGGGTA 815
QY 781 GGGGCTGTTGGTGGGACGTCGGGATCCTTAACCCCTCTGGACTCTCTGGGTGAGCTTCCTTA 840
Db 816 GGGGCTGTTGGTGGGACGTCGGGATCCTTAACCCCTCTGGACTCTCTGGGTGAGCTTCCTTA 875
QY 841 ACAGCCCTCGGATCTGGGTAGGCTCAATGTGGAGCGGGCTGTGACAGGCTACTTCCAC 900
Db 876 ACAGCCCTCGGATCTGGGTAGGCTCAATGTGGAGCGGGCTGTGACAGGCTACTTCCAC 935
QY 901 TGTGTATCTCGCATCAGCGGTTCTCTAGG 929
Db 936 TGTGTATCTCGCATCAGCGGTTCTCTAGG 964

RESULT 6
US-09-902-939-1
; Sequence 1, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: mouse
US-09-902-939-1

Query Match 63.3%; Score 678.2; DB 10; Length 1752;
Best Local Similarity 84.5%; Pred. No. 4.6e-159; Indels 18; Gaps 2;
Matches 793; Conservative 0; Mismatches 128

QY 10 CAGGACGGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGCTTGCCCGGAGGGTG 69
Db 7 CAGGACATGTCGAAGCCCTCAGAGGAGCGCTGTGCTCTGGCTTGCCAGGCGACA 66
QY 70 CTGACCATGCGCTTCGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTACGCCCGGG 129
Db 67 CTCAGATCATCTTTGGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTACGCCCGAGC 126
QY 130 GTGCGCTGGCTTCGATCGCTACGCGCTTCGCGCTTCGCGCTTACGGGCTTCCTT 189
Db 127 GTTCTCTGGCTTCAGATCGCTATGACTCTCTGGGCTTTGGGCTCTATGGGCACTTCCTC 186
QY 190 TCAGCGCACCTGGTGGCGAGAGCTCTTCGCGTACCTGGAGCACCGCGGGTGGCGGG 249
Db 187 AGCGCACACCTAGTGGCACAGAGCTCTTCGCTTACCTGGAGCACCGAAGGTGGCAGG 246
QY 250 GCGGCGCG-----GGGCGGCTGGATGACGACACCGCGCGCAGTGTGCGG 294
Db 247 GCTGCGCGCGCTCTCTTGGCGAAGGGGCGCCCTGGATGCGGCGCTACGCGAGCGTGGCA 306
QY 295 CTGACCATCTCGGCTTACGAGAGGACCCCGGCTACCTGCGGCTGCTGCGCTCGCGC 354
Db 307 CTCACCATCTCAGCCCTACCAAGAGATCCCGCTTACCTGCGGCTGCTTGAACCTCCGCG 366
QY 355 CGCGCCCTGCTGTACCGCGCGCGGCTGCGGCTCTCATGCTGCTGTGATGGCAACCGC 414
Db 367 CGCGCTTGTGTACCGCGCACAGAGGTTACGCGTCTCATGCTGCTGTGAGCGCAACCGC 426
QY 415 GCCGAGGACCTCTACATGCTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGC 474

Db 427 GCTGAGGATCTGTATACATGCTGGACATGTTCCGAGAAGTCTTCCCGATGAGGACCCCGC 486
QY 475 AGCTACGCTGTGGACCGCAACTACCAAGCCCTGGAAACCCCGCGCGG--CGGCGG 531
Db 487 ACTTATGTGTGGATGGCACTACCATCAGCCCTGGAAACCAAGCGGAGCTACGGCGCT 546
QY 532 GTGGCGCGCGGACCTATCGGAGGTGGAGCGGAGGATCTCTGGGCGGCTGGCAGTGGAG 591
Db 547 GTGCGTGAAGGTGCTTACCGGAGGTGGAGCGGAGGACCCCGCGGCTTGGCGGTGGAG 606
QY 592 GGGCTGTGAGGACTCGCAGGTGCTGCTGGGCGGAGCGCTGGGCGGCGCAAGCCCGAG 651
Db 607 GCGCTGTGAGAACACGACAGGTGCTGCTGGGCTCAGCGTTGGGGCGGCAAACTGAG 666
QY 652 GTCATGTACACAGCCCTCAAGGCGCTCGAGAGATTCGGTGGACTACGTGCGAGTCTGTGAC 711
Db 667 GTCATGTACACAGCTTCAAGGCACTGGGCGACTCGGTGGACTACGTGCGAGTCTGTGAC 726
QY 712 TCGGACACAAGGTGGACCCCATGTCCTGGAGCTCGTGGGCTACTGGACGAGGAC 771
Db 727 TCAGACACAAGACTAGACCCCATGTCCTGGAGCTTGTGCGAGTGTGGATGAAGAC 786
QY 772 CCGCGGTAGGGCTGTTGGTGGGAGCGTGGGATCTTTAACCTCTGAGACTCTCGGTC 831
Db 787 CCGCGGTAGGGCTGTTGGAGGGGATGTAGGATCTTAACTCTGAGCTCTCGGTC 846
QY 832 AGCTTCTTAAGCAGCGCTCGATACCTAGCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAG 891
Db 847 AGCTTCTTGAAGCAGCTTTCGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAG 906
QY 892 TACTTCCACTGTGTATCTTCGATCAGCGGTTCTCTAGGT 930
Db 907 TACTTCCACTGTGTCTCTGATCAGTGTCTCTGGGT 945

RESULT 7
US-10-363-345A-13975
; Sequence 13975, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 13975
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-Island No: 13975
US-10-363-345A-13975

Query Match 39.4%; Score 422; DB 18; Length 662;
Best Local Similarity 77.3%; Pred. No. 2.3e-95;
Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 34 GCAGCCCGCGCTGCTCGGCGCTGGCCGGAGGGTGTGACCATCGCTTCGCCCTGCTC 93
Db 1 GTAGTTTGTCTGTTGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTT 60
QY 94 ATCTTGGGCTCATGACCTGGGCTTACGCCCGCGGGTGGCGCTCGGCTCGGATCGCTAC 153
Db 61 ATTTGGGTTTATGATTTGGTTTACGCTCGGGGTGTGTTGGTTTGGTTTGGTTTGGTTT 120
QY 154 GCGCTCTGCGCTTGGGCTTACGGGCTTCTGCGGCGCTTCTTTCAGCGCACTCGTGGCGAGC 213

QY	95	TCCTGGGCTCATGACCTGGGCTTACGCGCGGGGTGCGCTTGGCTCCGATCGCTAGC	154
DB	601	TCCTAAACCTCATAACTAAACCTTACGCGCGGAAATACCGTAACTCCGATCGCTAGC	542
QY	155	GCCTCTCGGCTTCGGCTCTACGGGGCTTCTTTTCAGCGCACCTGGTGGCGGAGAGCC	214
DB	541	ACCTCTTAACCTTGACCTCTACGAAACCTTCTCTTTCAACGCACTTAATAACGCAAAACC	482
QY	215	TCTTTCGGGTACTTGGAGCACCGGGGGTGGGGCGCGCGCGGGGCGCTTGGATGCGAG	274
DB	481	TCTTTCGGGTACTTAAACACCGACGAATAACGACGACGCGGAAACCGCTAAATACAA	422
QY	275	CCACCGCGGAGTGTGGCGCTGACCATCTCGGCTTACAGAGAGACCCGCGTACCTGCG	334
DB	421	CCACCGCGGCAATATAACGCTAACCATCTCGGCTACCAAAAAACCCCGGTACCTAC	362
QY	335	GCCAGTGCCTGGCGTCCGCGCGGCTGCTGTATCCCGCGCGCGGCTGGCGTCTCTCA	394
DB	361	GCCANTACTTAACGTTCGCCCGGCCCTTACTATCCCGCGCGGCGACTACGCGTCTCTCA	302
QY	395	TGCTGGTGGATGGCAACCGCGCCGAGGACCTCTACATGGTGGACATGTTCCGCGAGGTCT	454
DB	301	TAATAATAATAATAACAACCGCGCGGAAACCTTACATAATCGACATATTCCGCGAAATCT	242
QY	455	TCGCTGACGAGGACCCGCGCACGTAAGTGTGGAGCGGCAACTACACAGACCTCGGGAAC	514
DB	241	TCGCTAACGAAAAACCCGCGCACGTACGTATATAAACGACAACTACCCACCAACCTTAAAAAC	182
QY	515	CCGCGGCGGGCGCGGTGGCGCGGAGCTTCCGGAGGTGGAGGCGGAGGATCCTG	574
DB	181	CCGCGACGACGACGCGAATAAACCGCGAAACCTATCGAAAAATAAAAAACGAAAAATCCTA	122
QY	575	GGCGGCTGGCAGTGGAGCGCTGGTGAGGACTCGCAGGTGGCTGTGGCGCAGCGCT	634
DB	121	AACGACTAACATATAAAACGCTATAAAAACTCGCAATACTGATACGTAACGCAACGCT	62
QY	635	GGGCGCGCAACGCGAGGTGATGATACACAGCTTCAAGCGCTCGGAGATTCGGTGGACT	694
DB	61	AAAAACGACAAACGCGAAATCATATACACAACTTCAAAACGCTCGAAAAATTCGATAAACT	2
QY	695	A 695	
DB	1	A 1	

RESULT 10
 US-10-363-345A-13974
 ; Sequence 13974, Application US/10363345A
 ; Publication No. US20040234960A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexander Olek
 ; APPLICANT: Christian Piepenbrock
 ; APPLICANT: Kurt Berlin
 ; TITLE OF INVENTION: Method for determining the degree of methylation of defined
 ; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3'
 ; FILE REFERENCE: E01/1227
 ; CURRENT APPLICATION NUMBER: US/10/363,345A
 ; CURRENT FILING DATE: 2003-03-03
 ; NUMBER OF SEQ ID NOS: 40712
 ; SEQ ID NO 13974
 ; LENGTH: 662
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; OTHER INFORMATION: CpG-island No: 13974
 US-10-363-345A-13974

Query Match	36.8%	Score 393.8	DB 18	Length 662
Best Local Similarity	74.7%	Pred. No. 2.4e-88		
Matches 494	Conservative 0	Mismatches 167	Indels 0	Gaps 0

QY 35 CAGCCCGCCGCTGCTCCGGCTCGCCCGGAGGGTGTCTGACCATCGCCCTTCGCCCTGCTCTCA 94

Db	2	CAACCTACCGGTACTCCGACCTAACCCGAAAAATACTAACCATCGCCTTCGCGCCTACTCA	61
Qy	95	TCCTGGGCGCTCATGACCTGGGCGCTACGCGCGCGGGTGCCTGCGCTCCGATCCGATCGCTAG	154
Db	62	TCCTAAACCTCTAAACCTTAACCTTACGCGCGCGAAATACCGCTAACCTCCGATCGCTAG	121
Qy	155	GCCTCTGGGCTTCGGGCTCTACGGGGCGTTCCTTTTCAGCGCACTGTGTGGCGCAGAGCC	214
Db	122	ACCTCTCTAAACCTTCGACCTCTACGAAACCTTCTTCAACGCACTTAATACGCAAAAC	181
Qy	215	TCCTCGGCTACCTGGAGCACCGGCGGGTGGCGGGCGCGCGGGGGCGCTGGATGCAG	274
Db	182	TCCTTCGGGTACCTTAAACACCGACGAAATAACGACGACGCGGAAAAACCGCTAAATACAA	241
Qy	275	CCACCGCGCGCAGTGTGGCGCTGACCATCTCCGGCTCACGAGGAGGACCCCGGTAACCTGC	334
Db	242	CCACCGCGCGCAATATAACGCTAACCATCTCCGGCTACCAAAAAAACCCCGGTACCTAC	301
Qy	335	GCCTGCTGGGCTGCGCGCGCGCGCTGTGTATCCCGCGCGCGCGGCTGCGCGTCTCTCA	394
Db	302	GCCAATACCTAAACGTCGCGCGCGCGCTTACTATACCCGCGCGCGACTACGCGTCTCTCA	361
Qy	395	TGTTGGTGGATGGCAACCGCGCGAGACCTCTACATGTCGACATGTTCCGCGAGGCTCT	454
Db	362	TAATAATAATAACACCGCGCGGAAACCTCTCATTAATCGACATATTCGCGAAATCT	421
Qy	455	TCGCTGACGAGGACCCCGCCACGCTACGTGTGGGACGGCAACTACCAACGACCTCGGGAC	514
Db	422	TCGCTAACGAAACCCCGCCACGTACGTATAAAACGACCAACTACCAACCAACCTAAAAAC	481
Qy	515	CCGCGCGCGCGCGCGGTGGCGCGGAGCTATCGGGAGGTGGAGCGGAGGATCCTTG	574
Db	482	CCGCGACGAGAACGCGATAAACCCCGGAAACCTATCGAAAAATAAAAACGAAAAATCCTTA	541
Qy	575	GGCGGCTGGCAGTGGAGCGCTGCTGAGGACTCCGACGTGCGTGTGGCGCGACGCT	634
Db	542	AACGACTAACATAAAAAACGCTATAAAAACTCGCAATACGTATACGTAAACGCAACGCT	601
Qy	635	GGGCGCGCAAGCGAGGTCAATGACAGCGCTTCAAGCGCTCGAGATTTCGGTGGACT	694
Db	602	AAAAACGAAACCGGAAATCATATACACAACTTCMAAACGCTCGAAAAATTCGATAAACT	661
Qy	695	A 595	
Db	662	A 662	

RESULT 11

US-09-918-995-27210

; Sequence 27210, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 27210

; LENGTH: 490

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(490)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-27210

```

RESULT 11
US-09-918-995-27210
; Sequence 27210, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27210

```



```
RESULT 13
US-10-188-832-136
; Sequence 136, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Azi, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 136
; TYPE: DNA
; LENGTH: 1662
; ORGANISM: Homo sapiens
; US-10-188-832-136

Query Match      25.5%; Score 272.8; DB 17; Length 1662;
Best Local Similarity 61.0%; Pred. No. 2.6e-58; Indels 39; Gaps 4;
Matches 533; Conservative 0; Mismatches 302;

QY 61 CGAGGGTGCTGACCATCGCTTCGCGCTGCTCATCTCGGGCTCATGACCTGGGCGCTAC 120
DB 28 CGTGTGTGGGACACGAGCTGTTGCCCTGGCAGTGCTGGGTGGCATCTCGCAGCTAT 87
QY 121 GCGCGGGGTGCGGTGCGCTCGATCGCTACGGGCTCTGGGCTTGGGCTCTACGGG 180
DB 98 GTGAGGGGTACAGTTTATCCACACGAGAAAGACATACCTGTCTCTCGGCTGTACGGC 147
QY 181 GCGTTCTTTACGACACCTGTGGCGAGAGCTCTTGGCTTCTGAGCAGCAGCGCGG 240
DB 148 GCATCTCTGGGCTGACCTGCTCATTTAGAGCCCTTTTGGCTTCTTGGAGCAGCGGCG 207
QY 241 GTGGCGGGCGGCGGCGGCGCTGTGATGAGCAGCAGCGCGC---GCAAGTGTGGCGGTG 297
DB 208 ATGCGAGCTGCGCGGCGAGCCCTGAAGCTGCCCTCCCGCGCGGGGCTCGGTGGCACTG 267
QY 298 ACCATCTCGCTACAGAGAGACCCCGGTACCTGCGCCAGTGTGCGGCTGCGGCGCGC 357
DB 268 TGCAATTGCGGCTACAGAGAGACCTGACTACTTGGCAAGTGTGCTGCGCTCGGCGCAG 327
QY 358 GCGCTGTGTACCGGCGGCGGCTGCGGCTCTCATGCTGTGATGGTGGAGCAACCGCGC 417
DB 328 CGCATCTCTCTCCCTGAC-----CTCAAGGTGGTGTGATGGTGGAGCAACCGCGCAG 381
QY 418 GAGGACCTCTACATGCTGCATGTTTCCGCGAGGTCTTCGCTGAGGAGACCCCGCAGG 477
DB 382 GAGGAGCGCTACATGCTGCATCTTCCAGAGGTGTGCTGGGCGGCGACCGAGCAGCGCGC 441
QY 478 TACGTGTGGGACGGCAACTACACAGCCCTTGGGAACCCCGCGCGCGGCGCGGTGGGC 537
DB 442 TTTCTTTGTGGCGCAGCAACTTCCA-----TGAGGAGCGGAGGTGAG 486
QY 538 GCGGAGCGCTATCGGAGGTGAGGCGGAGGATCTTGGCGCGCTGGCAGTGGAGCGGTG 597
DB 487 ACGGAGCGCAGCTCGAGGAGGCGATGACCGT-----GTGCGGATGTG 531

RESULT 14
US-10-309-560-9
; Sequence 9, Application US/10309560
; Publication No. US20030235893A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: KUMARI, KSHAMA
; TITLE OF INVENTION: HYALURONAN SYNTHASES AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 35541.082
; CURRENT APPLICATION NUMBER: US/10/309,560
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/336,105
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Xenopus laevis
; US-10-309-560-9

Query Match      23.2%; Score 248.4; DB 17; Length 1767;
Best Local Similarity 56.3%; Pred. No. 3e-52;
Matches 490; Conservative 0; Mismatches 371; Indels 9; Gaps 1;

QY 62 GGAGGGTCTGACCATCGCTTCGCGCTGCTCATCTCGGCTCATGACCTGGGCTAGC 121
DB 80 GGAGGATTAATTATTATTCTTTTGTGTGGTGTATTAGCTACCATTAACAGCAGCTATG 139
QY 122 CGCGGGGTGCGGTGCGCTCCGATCGCTACGGCTCTCGGCTCTCGGCTCTCGGCTCTACGGGG 181
DB 140 TGGCAGAGTTCAGGTCTCTCAACATGAAGCAATTTCTTCTCTCTCGGCTTTATGTC 199
QY 182 CTTCTCTTTACGCGACCTGTGGCGCAGAGCTCTTCGCGTACCTGAGACACCGCGGG 241
DB 200 TTGCAATCTTCTCCACCTGATGACAGAGCTCTTTTGGCTTCTGAGATACGAGG 259
QY 242 TGGCGGCGGCGCGGGGCGCTGGATGAGCCACCGCGCAGTGTGGGCTGACCA 301
DB 260 TAAATAAGAGTGA-----GCTTCTTGCAGCTTTTAAAGAACAGTGGCTCTGACCA 310
QY 302 TCTCGGCTTACAGGAGACCCCGGTACCTGCGCAGTGTGCGCTCGGCTCGGCGCGCC 361
DB 311 TTGCTGGGTATCAGGAGAACCTTGAGTACCTGATAAAGTGTGGAATCTCTCAAGTATG 370
QY 362 TGCTGTATCCCGCGCGCGGTGCGGCTCTCATGTGTGTGATGGCAACCGCGCGGAGG 421
DB 371 TGAATAACCCCAAGATAAACTCAAGATCAATTTTGGTTCATCGATGGGAACACAGAGATG 430
```

Qy	422	ACCTCTACATGTCGACATGTTCTCCGAGGTC	TTGCTGACGAGGACCCCGCACGTCACG	481
Db	431	ATGCCCTACATGATGGAGATGTTTCAAAGACG	TGTTCCACGCTGAAGATGTAGGCACCTACG	490
Qy	482	TGTGGGACCGCAACTACCAACAGCCTCGGGA	ACCGCGGGCGCGGTGGGGCGCGG	541
Db	491	TATGGAGGGAAATTACCACACTGTTAAAAAG	CCCTGAGGAGACCAATAAGGGATCTCTGTC	550
Qy	542	GAGCCTATCCGGAGGTGGAGGCGGAGGATCC	TGGCGCGCTGGCAGTGGAGCGCTCGTGA	601
Db	551	CTGAGGTTTCTAAGCCCTTGAATGAAGATGA	AGGTATCAATATGCTGGAAGAACTTGTTA	610
Qy	602	GGNCTCGAAGTTCGTGTGCTGGCGCGACCC	CTGGGCGCGCAGCGCGAGGTCAATGACA	661
Db	611	GAAACAAGAGATGTGTGTGCATCATGCAACA	GAGTGGGGCGGAAAGAGAGTCAATGACA	670
Qy	662	CAGCCTTCAAGCGCGTCGGAGATTCGGTGG	ACTACTGTCAGGTCCTGTGACTTCGGACACA	721
Db	671	CAGCATTCACGCGCATTTGGGACTTCTGTG	AGCTATGTACAGGTCCTGTGACTTCGGACACA	730
Qy	722	GGTTGACCCCATGCGACTGCTGGAGTCTGT	CGGGTACTGGACGAGGACCCCGGGTAG	781
Db	731	AACTGGATGAACCTGGCAACAGTGGAAAT	TGCTGAAGGTTCTGGAATCCAATGACATGACG	790
Qy	782	GGGCTGTGTGGTGGGACGTGGGATCCCTTA	ACCCTCTGGACTCTCTGGGTGAGCTTCTTA	841
Db	791	GGCGAGTGGGAGGACGTTGCGATTCCTGA	ACCCTTATGATTTCTTCATTAGTTTTCATGA	850
Qy	842	GCAGCTCGCGATCTCTGGGTAGCCTTCAAT	GTGGAGCGGGCTTGTACAGAGCTACTTCCACT	901
Db	851	GCAGCTCGGTTACTGGAATGGGTTTAACT	GTGGAGGAGGCGCTGCCAGTCTTACTTCGACT	910
Qy	902	GTGTATCTGCATCAGCGGTTCTCTAGGTA		931
Db	911	GGGTGTCCTGTATAAGTGGACCTCTGGGAA		940

```

RESULT 15
US-09-902-939-3
; Sequence 3, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHaza
; APPLICANT: Weiliam Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-US0
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: mouse
US-09-902-939-3

```

Qy	241	GTGCGGCGGCGGCGG-----GGGCGCGTGTGATGAGCCACCGCGCGCAGTGTGGCG	294
Db	208	ATGCGCAGGCGAGGCGGCGCCCTCAAGCTGCACCTGCTCCAGAGGTCTCGGTTCAGTGGCA	267
Qy	295	CTGACCATCTCGGCTACCAAGAGGACCCCGGTACTGCGCGCAGTGTGCTGGCGTCCGCG	354
Db	268	CTCTGCATTTGTGCTTACCAAGAGGACCCCGAATACCTGCGCAAGTGCCTTCGCTCAGCT	327
Qy	355	CGCGCCCTGTGTATACC CGCGCGCGCGGTGCGCGTCTTCATGTGTGGTGGATGGCAACGCG	414
Db	328	CAGGCGCATTGCCTTTCCAAAC-----CTCAAGGTGGTCACTGAGTGGATGGCAATGCG	381
Qy	415	GCCGAGGACCTTACATGGTGCACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCG	474
Db	382	CAGGAAGATACCTACATGTTGGACATCTTTCATGAGGTGTGGTGGCACTGAGCAAGCT	441
Qy	475	ACGTACGTGTGGGACGGCAACTCACGAGCCCTGGGAAACCGCGCGCGCGCGCGGTG	534
Db	442	GGCTTCTTTGTGTGGCGTAGCAATTTCCA-----TGAGGCGGGTGAAGGA	486
Qy	535	GGCGCCGAGCCTATCGGAGGTGGAGCGGAGGATCCTGGGCGGCTGGCAGTGGAGCG	594
Db	487	GAGACAGAGGCCAGCCTGCGAAGGAGGCAATGAGCGT-----GTGCGAGCT	531
Qy	595	CTGTGTGAGGAATCGCAGGTGCGTGTGGGTGCGCAGCGCTGGGGCGGCAAGCGCGAGGTC	654
Db	532	GTGTGTGGGCGACCACTTCTCATGCATCATGCAGAAGTGGGGGGCAAGCGTCAGGTC	591
Qy	655	ATGTACACAGCCTTCAAGGCGCTCGGAGATTCGGTGTGACTACGTGCAGGTCTGTGACTCG	714
Db	592	ATGTACACTGCTCTTCAAGGCGCTTTGGCAACTCAGTGGACTACATCCAGGTGTGTACTCT	651
Qy	715	GACACAGGTTGGACCCCATGGCACCTGTGAGCTCGTGGCGGTACTGGACGAGACCCC	774
Db	652	GACACTGTGTGGACCCAGCCTGCACCATTGAGATGCTTCAAGTCTTGGAGAAGATGCC	711
Qy	775	CGGGTAGGGCTGTTGTGTGGGACGTGCGGATCTTAAACCTCTGGACTCTCTGGGTACAG	834
Db	712	CAAGTAGGAGGTGTGGAGGAGATGTCCAAATCTCCAAAGATATGATTCATGGATCTCC	771
Qy	835	TTCCTAAGCAGCCGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTTCAGAGCTAC	894
Db	772	TTCTCTGAGCAGTGTGAGGTACTGATGGCTTTCAACGTGGAGCGGCGCTGCCAGTCTTAC	831
Qy	895	TTCCACTGTGTATCCTGCATCAGCGGTTCTCTAGGTA	931
Db	832	TTTGGCTGTGTGCAATGTATTAGTGGGCTTTGGGCA	868

Search completed: March 13, 2005, 17:09:55
Job time : 758.899 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:01:16 ; Search time 210.96 Seconds
(without alignments)
8307.054 Million cell updates/sec

Title: US-10-672-399-3
Perfect score: 1071
Sequence: 1 atgagacagcaggacgccc.....ccggaccatcggtgatga 1071

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	927.4	86.6	2116	3	US-08-635-552A-1
2	924.2	86.3	2087	4	US-09-949-016-5659
3	915	85.4	2117	3	US-09-155-768-1
4	697.2	65.1	14862	4	US-09-949-016-17401
5	396.4	37.0	601	4	US-09-949-016-201775
6	272.8	25.5	1662	4	US-09-949-016-3728
7	236.2	22.1	1665	4	US-08-812-008-31
8	230.8	21.5	601	4	US-09-949-016-201780
9	176.2	16.5	2947	4	US-08-675-499A-1
10	176.2	16.5	2947	4	US-08-812-008-1
11	167.6	15.6	9871	4	US-09-949-016-15470
12	148.8	13.9	3003	4	US-09-949-016-737
13	137.4	12.8	601	4	US-09-949-016-132199
14	82.8	7.7	6645	4	US-09-902-540-8819
15	82.8	7.7	10210	4	US-09-902-540-938
16	69.2	6.5	1320	4	US-09-724-797-85
17	66.6	6.2	3957	4	US-10-237-551-193
18	66.6	6.2	154746	4	US-09-827-688-8
19	66.6	6.2	154746	4	US-09-827-688-8
20	66.2	6.2	1614	4	US-09-616-289-45
21	66.2	6.2	12425	4	US-09-616-289-50
22	62.2	5.8	32176	4	US-09-949-016-12479
23	62	5.8	1292	3	US-08-483-533-37
24	62	5.8	1292	3	US-09-283-471A-37
25	60.6	5.7	4257	2	US-08-690-473-1
26	60.6	5.7	4257	3	US-09-259-821A-1
27	60.6	5.7	4257	3	US-08-843-659-1

28	60.6	5.7	4257	4	US-09-825-288A-1	Sequence 1, Appl
29	60.6	5.7	12001	1	US-08-458-568A-11	Sequence 11, Appl
30	60.2	5.6	2900	1	US-08-034-650-9	Sequence 9, Appl
31	60.2	5.6	2900	1	US-08-449-015-9	Sequence 9, Appl
32	59.6	5.6	1647	4	US-09-902-540-4211	Sequence 4211, Ap
33	59.6	5.6	21330	4	US-09-902-540-1209	Sequence 1209, Ap
34	59.4	5.5	1926	3	US-09-249-585A-4	Sequence 4, Appl
35	59.4	5.5	1931	2	US-09-130-114-2	Sequence 2, Appl
36	59.4	5.5	441529	3	US-09-103-840A-1	Sequence 1, Appl
37	59.2	5.5	1051	2	US-08-865-273-1	Sequence 1, Appl
38	59.2	5.5	1051	3	US-09-385-174-1	Sequence 1, Appl
39	59.2	5.5	1182	4	US-09-673-198-31	Sequence 31, Appl
40	59.2	5.5	1361	4	US-09-614-912-37	Sequence 37, Appl
41	59	5.5	4403765	3	US-09-103-840A-2	Sequence 2, Appl
42	58.4	5.5	5808	4	US-09-902-540-794	Sequence 794, App
43	58.2	5.4	967	4	US-09-902-540-7886	Sequence 7886, Ap
44	57.8	5.4	8604	4	US-09-902-540-5331	Sequence 5331, Ap
45	57.8	5.4	33399	4	US-09-902-540-1260	Sequence 1260, Ap

ALIGNMENTS

RESULT 1
US-08-635-552A-1
; Sequence 1, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
US-08-635-552A-1

Query Match 86.6%; Score 927.4; DB 3; Length 2116;
Best Local Similarity 99.9%; Pred No. 3.2e-175;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGACAGCAGGACGCCCAAGCCCACTCCTGACGCCCGCTGCTCGGCTGGCC 60
|||||

```
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCCGGCTCGGCC 95
Qy 61 CGAGGGTGTGACCATCGCCTTGCCCTGTCTATCTTGGGCTCATGACCTTGGGCTTAC 120
Db 96 CGAGGGTGTGACCATCGCCTTGCCCTGTCTATCTTGGGCTCATGACCTTGGGCTTAC 155
Qy 121 GCGCGCGGGTGGCGCTGGCTCGATCGCTACCGGCTCTTGGGCTTGGGCTTACGGG 180
Db 156 GCGCGCGGGTGGCGCTGGCTCGATCGCTACCGGCTCTTGGGCTTGGGCTTACGGG 215
Qy 181 GCCTTCCTTTTTCAGCGCACCTGTGTGCGCAGAGCCTCTTCCGGTACTTGGAGCACCGGG 240
Db 216 GCCTTCCTTTTTCAGCGCACCTGTGTGCGCAGAGCCTCTTCCGGTACTTGGAGCACCGGG 275
Qy 241 GTGGCGCGCGCGCGCGGGGCGCTGTGATGACGCCACCGCGCGCAGTGTGGCGCTTAC 300
Db 276 GTGGCGCGCGCGCGCGGGGCGCTGTGATGACGCCACCGCGCGCAGTGTGGCGCTTAC 335
Qy 301 ATCTCCGCTTACCAAGGACCCCGGCGCTGTGATGACGCCACCGCGCGCAGTGTGGCGCTTAC 360
Db 336 ATCTCCGCTTACCAAGGACCCCGGCGCTGTGATGACGCCACCGCGCGCAGTGTGGCGCTTAC 395
Qy 361 CTGCTGTACCGCGCGCGCGCGCGCTGCGCTCTCTATGTGTGGATGGCAACCGCGCGCGAG 420
Db 396 CTGCTGTACCGCGCGCGCGCGCGCTGCGCTCTCTATGTGTGGATGGCAACCGCGCGCGAG 455
Qy 421 GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTCCGTCGACGAGGACCCCGCCACGTTAC 480
Db 456 GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTCCGTCGACGAGGACCCCGCCACGTTAC 515
Qy 481 GTGTGGACCGGCAACTACCAACCGCTGGGAAACCCCGCGCGCGCGCGGCTGGGCGCC 540
Db 516 GTGTGGACCGGCAACTACCAACCGCTGGGAAACCCCGCGCGCGCGCGGCTGGGCGCC 575
Qy 541 GTGTGGACCGGCAACTACCAACCGCTGGGAAACCCCGCGCGCGCGCGGCTGGGCGCC 600
Db 576 GTGTGGACCGGCAACTACCAACCGCTGGGAAACCCCGCGCGCGCGCGGCTGGGCGCC 635
Qy 601 AGGACTCGCAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 660
Db 636 AGGACTCGCAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 695
Qy 661 ACAGCCTTCAAGCGGCTCGGAGATTCGATGGAATTCGATGGAATTCGATGGAATTCGATG 720
Db 696 ACAGCCTTCAAGCGGCTCGGAGATTCGATGGAATTCGATGGAATTCGATGGAATTCGATG 755
Qy 721 AGGTTGGACCCCATGGCAGCTGTGGAGCTGTGGAGCTGTGGAGCTGTGGAGCTGTGGAG 780
Db 756 AGGTTGGACCCCATGGCAGCTGTGGAGCTGTGGAGCTGTGGAGCTGTGGAGCTGTGGAG 815
Qy 781 GGGGCTGTGTGGGAGCTGCGGATCTTAAACCTCTGGAATCTGCGGTGAGCTTCTTCTTA 840
Db 816 GGGGCTGTGTGGGAGCTGCGGATCTTAAACCTCTGGAATCTGCGGTGAGCTTCTTCTTA 875
Qy 841 AGCAGCTGCGCATCTGGGTAGCTTCAATCTGAGCGGGCTTCTGAGCTACTTCCAC 900
Db 876 AGCAGCTGCGCATCTGGGTAGCTTCAATCTGAGCGGGCTTCTGAGCTACTTCCAC 935
Qy 901 TGTGTATCTTCATCAGCGGTTCTCTAGG 929
Db 936 TGTGTATCTTCATCAGCGGTTCTCTAGG 964
```

RESULT 2

US-09-949-016-5659
; Sequence 5659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5659

Query Match 86.3%; Score 924.2; DB 4; Length 2087;
Best Local Similarity 99.7%; Pred. No. 1.4e-174;
Matches 926; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCCGGCTCGGCC 60
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCCGGCTCGGCC 95
Qy 61 CGAGGGTGTGACCATCGCCTTGCCCTGTCTATCTTGGGCTCATGACCTTGGGCTTAC 120
Db 96 CGAGGGTGTGACCATCGCCTTGCCCTGTCTATCTTGGGCTCATGACCTTGGGCTTAC 155
Qy 121 GCGCGCGGGTGGCGCTGGCTCGATCGCTACCGGCTCTTGGGCTTGGGCTTACGGG 180
Db 156 GCGCGCGGGTGGCGCTGGCTCGATCGCTACCGGCTCTTGGGCTTGGGCTTACGGG 215
Qy 181 GCCTTCCTTTTTCAGCGCACCTGTGTGCGCAGAGCCTCTTCCGGTACTTGGAGCACCGGG 240
Db 216 GCCTTCCTTTTTCAGCGCACCTGTGTGCGCAGAGCCTCTTCCGGTACTTGGAGCACCGGG 275
Qy 241 GTGGCGCGCGCGCGCGGGGCGCTGTGATGACGCCACCGCGCGCAGTGTGGCGCTTAC 300
Db 276 GTGGCGCGCGCGCGCGGGGCGCTGTGATGACGCCACCGCGCGCAGTGTGGCGCTTAC 335
Qy 301 ATCTCCGCTTACCAAGGACCCCGGCGCTGTGATGACGCCACCGCGCGCAGTGTGGCGCTTAC 360
Db 336 ATCTCCGCTTACCAAGGACCCCGGCGCTGTGATGACGCCACCGCGCGCAGTGTGGCGCTTAC 395
Qy 361 CTGCTGTACCGCGCGCGCGCGCTGCGCTCTCTATGTGTGGATGGCAACCGCGCGCGAG 420
Db 396 CTGCTGTACCGCGCGCGCGCGCTGCGCTCTCTATGTGTGGATGGCAACCGCGCGCGAG 455
Qy 421 GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTCCGTCGACGAGGACCCCGCCACGTTAC 480
Db 456 GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTCCGTCGACGAGGACCCCGCCACGTTAC 515
Qy 481 GTGTGGACCGGCAACTACCAACCGCTGGGAAACCCCGCGCGCGCGGCTGGGCGCC 540
Db 516 GTGTGGACCGGCAACTACCAACCGCTGGGAAACCCCGCGCGCGCGGCTGGGCGCC 575
Qy 541 GAGCCTTATCGGAGGTGAGGAGGAGATCTTGGGCGGCTTGGCAGTGGAGGCGCTGGTG 600
Db 576 GAGCCTTATCGGAGGTGAGGAGGAGATCTTGGGCGGCTTGGCAGTGGAGGCGCTGGTG 635
Qy 601 AGGACTCGCAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 660
Db 636 AGGACTCGCAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 695
Qy 661 ACAGCCTTCAAGCGGCTCGGAGATTCGATGGAATTCGATGGAATTCGATGGAATTCGATG 720
Db 696 ACAGCCTTCAAGCGGCTCGGAGATTCGATGGAATTCGATGGAATTCGATGGAATTCGATG 755
Qy 721 AGGTTGGACCCCATGGCAGCTGTGGAGCTGTGGAGCTGTGGAGCTGTGGAGCTGTGGAG 780
Db 756 AGGTTGGACCCCATGGCAGCTGTGGAGCTGTGGAGCTGTGGAGCTGTGGAGCTGTGGAG 815
Qy 781 GGGGCTGTGTGGGAGCTGCGGATCTTAAACCTCTGGAATCTGCGGTGAGCTTCTTCTTA 840
```


Db 816 GGGGCTGTGGTGGGATGTGGGATCTTAACCTCTTGGACTCTGGACTCTGGGTGAGCTTCCTA 875
Qy 841 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTTACTTCCAC 900
Db 876 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTTACTTCCAC 935
Qy 901 TGTGTATCTGCATCAGCGGTCTCTAGG 929
Db 936 TGTGTATCTGCATCAGCGGTCTCTAGG 964

RESULT 3

US-09-155-768-1

; Sequence 1, Application US/09155768A

; Patent No. 6162908

; GENERAL INFORMATION:

; APPLICANT: SEIKAGAKU CORPORATION

; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF

; FILE OF INVENTION: HUMAN-ORIGIN AND DNA ENCODING THE SAME

; FILE REFERENCE: TOYAM37.001APC

; CURRENT APPLICATION NUMBER: US/09/155,768A

; EARLIER FILING DATE: 1998-10-05

; EARLIER APPLICATION NUMBER: JP 8-084326

; EARLIER FILING DATE: 1996-04-05

; EARLIER APPLICATION NUMBER: JP 8-109663

; EARLIER FILING DATE: 1996-04-30

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 2117

; TYPE: DNA

; ORGANISM: HUMAN

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (149)...(1777)

US-09-155-768-1

Query Match

Best Local Similarity 85.4%; Score 915; DB 3; Length 2117;

Matches 918; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 7 CAGCAGGACGCGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGGCTGGCCCGGAGG 66
Db 50 CAGCAGGACGCGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGGCTGGCCCGGAGG 109
Qy 67 GTGCTGACCATCGCTTCGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTACGCCGCC 126
Db 110 GTGCTGACCATCGCTTCGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTACGCCGCC 169
Qy 127 GGGGTGCGGCTGGCTTCGGATCGCTACGCGCTCTCTGGGCTTCTAGGGGCTTC 186
Db 170 GGGGTGCGGCTGGCTTCGGATCGCTACGCGCTCTCTGGGCTTCTAGGGGCTTC 229
Qy 187 CTTTCAGGCGACCTGGTGGCGAGAGCTTTCGGGTACTGGAGACCGCGGGGTGGG 246
Db 230 CTTTCAGGCGACCTGGTGGCGAGAGCTTTCGGGTACTGGAGACCGCGGGGTGGG 289
Qy 247 GCGGGCGCGCGGGCGCGCTGATCAGCAGCCAGCGCGAGTGTGGCTGACCATCTCC 306
Db 290 GCGGGCGCGCGGGCGCGCTGATCAGCAGCCAGCGCGAGTGTGGCTGACCATCTCC 349
Qy 307 GCCTACCAAGGAGACCGCGGCTTACCTGGCGAGTGTGGGCTTCGCCCGCGGCTTCTG 366
Db 350 GCCTACCAAGGAGACCGCGGCTTACCTGGCGAGTGTGGGCTTCGCCCGCGGCTTCTG 409
Qy 367 TACCGCGCGCGGCTGGGCTTCCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 426
Db 410 TACCGCGCGCGGCTGGGCTTCCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 469
Qy 427 TACATGT 486
Db 470 TACATGT 529

Qy 487 GACGCAACTACCACTAGCCCTTGGGAACCCCGCGCGCGCGCTGGGCTGGGCGCCGAGACC 546
Db 530 GACGCAACTACCACTAGCCCTTGGGAACCCCGCGCGCGCGCTGGGCTGGGCGCCGAGACC 589
Qy 547 TATCGGAGGTGGAGGCGGAGATCCTGGGCGGCTGGCAGTGGAGCGGCTGTGAGGACT 606
Db 590 TATCGGAGGTGGAGGCGGAGATCCTGGGCGGCTGGCAGTGGAGCGGCTGTGAGGACT 649
Qy 607 GCGAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTCTATGTACACAGCC 666
Db 650 GCGAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTCTATGTACACAGCC 709
Qy 667 TTCAAGCGCTCGGAGATTCGGTGAATCTGTCAGTGTGTGACTCGGACACAAAGTTG 726
Db 710 TTCAAGCGCTCGGAGATTCGGTGAATCTGTCAGTGTGTGACTCGGACACAAAGTTG 769
Qy 727 GACCCCATGGCACTGCTGGAGCTCGTGGGCTACTGGAGCAGGAGCCCGGGGTAGGGCT 786
Db 770 GACCCCATGGCACTGCTGGAGCTCGTGGGCTACTGGAGCAGGAGCCCGGGGTAGGGCT 829
Qy 787 GTTGTGGGAGCTGCGGATCCTTAACCTCTGGACTCTGGGTGAGCTTCTTAAGCAGC 846
Db 830 GTTGTGGGAGCTGCGGATCCTTAACCTCTGGACTCTGGGTGAGCTTCTTAAGCAGC 889
Qy 847 CTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCACTGTGA 906
Db 890 CTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCACTGTGA 949
Qy 907 TCCTGCATCAGCGGTCTCTAGG 929
Db 950 TCCTGCATCAGCGGTCTCTAGG 972

RESULT 4

US-09-949-016-17401

; Sequence 17401, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: C1001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17401

; LENGTH: 14862

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-17401

Query Match

Best Local Similarity 65.1%; Score 697.2; DB 4; Length 14862;

Matches 699; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 CAGCAGGACGCGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGGCTGGCCCGGAGG 66
Db 6073 CAGCAGGACGCGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGGCTGGCCCGGAGG 6132
Qy 67 GTGCTGACCATCGGCTTTCGGCTTCTGCTCATCTCTGGGCTTCTGAGCTGAGGCTTTC 126
Db 6133 GTGCTGACCATCGGCTTTCGGCTTCTGCTCATCTCTGGGCTTCTGAGCTGAGGCTTTC 6192
Qy 127 GGGGTGCGGCTGGGCTTCGGATCGCTAGGCTTCTGGGCTTTCGGGCTTCTAGGGGCTTC 186
Db 6193 GGGGTGCGGCTGGGCTTCGGATCGCTAGGCTTCTGGGCTTTCGGGCTTCTAGGGGCTTC 6252

```
QY 187 CTTTCAGCGCACCTGGTGGCGAGAGCTCTTCGGGTACCTGGAGCACCGCGGGTGGCG 246
Db 6253 CTTTCAGCGCACCTGGTGGCGAGAGCTCTTCGGGTACCTGGAGCACCGCGGGTGGCG 6312
QY 247 GCGGGCGCGCGGGGCGCGCTGATGACGACACCGCGCAGTGTGGCGCTGACCATCTCC 306
Db 6313 GCGGGCGCGCGGGGCGCGCTGATGACGACACCGCGCAGTGTGGCGCTGACCATCTCC 6372
QY 307 GCCTACCAAGGAGACCCCGCGTACTCGCGCAGTGCCTGGCGTCCGCGCGCCCTGCTG 366
Db 6373 GCCTACCAAGGAGACCCCGCGTACTCGCGCAGTGCCTGGCGTCCGCGCGCCCTGCTG 6432
QY 367 TACCGCGCGCGCGCTCGCGTCTCTATGTGTGGATGCAACCGCGCGCGAGACCTC 426
Db 6433 TACCGCGCGCGCGCTCGCGTCTCTATGTGTGGATGCAACCGCGCGCGAGACCTC 6492
QY 427 TACATGTCGACATGTTCCCGGAGGTCTTCGCTGACGAGGACCCCGCCACGTACTGTGG 486
Db 6493 TACATGTCGACATGTTCCCGGAGGTCTTCGCTGACGAGGACCCCGCCACGTACTGTGG 6552
QY 487 GACGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGGTGGCGCCCGAGCC 546
Db 6553 GACGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGGTGGCGCCCGAGCC 6612
QY 547 TATCGGAGGTGGAGCGGAGATCTTGGCGGCTGGCAGTGGAGGCGCTGTGAGACT 606
Db 6613 TATCGGAGGTGGAGCGGAGATCTTGGCGGCTGGCAGTGGAGGCGCTGTGAGACT 6672
QY 607 CGCAGTGCCTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTCTATGTACACAGCC 666
Db 6673 CGCAGTGCCTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTCTATGTACACAGCC 6732
QY 667 TTCAGGCGCTCGGAGATTTCGGTGGACTACGTGCAGGTCTGT 708
Db 6733 TTCAGGCGCTCGGAGATTTCGGTGGACTACGTGCAGGTGAGT 6774

RESULT 5
US-09-949-016-201775/c
; Sequence 201775, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201775
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201775

Query Match 37.0%; Score 396.4; DB 4; Length 601;
Best Local Similarity 99.2%; Pred. No. 6.6e-70;
Matches 397; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 309 CTACAGAGAGACCCCGGTACTTGGCGCAGTGGCTGGCGTCCGCGCGCCCTGCTGTA 368
Db 601 CTACAGAGAGACCCCGGTACTTGGCGCAGTGGCTGGCGTCCGCGCGCCCTGCTGTA 542
QY 369 CCGCGCGCGCGCTCGCGCTCTCATGTGTGGATGGCAACCGCGCGAGGACCTCTA 428
|
```

```
Db 541 CCGCGCGCGCGCTGGCGTCTCTCATGTGTGGTGAACCGCGCGGAGGACCTCTA 482
QY 429 CATGGTGCACATGTTCCGCGAGGTCTTCGTGACGAGAGACCCCGCCACGTACGTGTGGGA 488
|
|
|
Db 481 CATGGTGCACATGTTCCGCGAGGTCTTCGTGACGAGAGACCCCGCCACGTACGTGTGGGA 422
QY 489 CGGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGTGGCGCGCGAGCCTA 548
|
|
|
Db 421 CGGCACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGTGGCGCGCGAGCCTA 362
QY 549 TCGGAGGTGAGAGCGGAGGATCTTGGGCGGCTGGCAGTGGAGCGCTGTGTGAGACTCG 608
|
|
|
Db 361 TCGGAGGTGAGAGCGGAGGATCTTGGGCGGCTGGCAGTGGAGCGCTGTGTGAGACTCG 302
QY 609 CAGTGTGCTGTGCTGGCGCAGCGCTGGGCGGCAAGCGGAGGTCTATGTACAGACCTT 668
|
|
|
Db 301 YAGTGTGCTGTGCTGGCGCAGCGCTGGGCGGCAAGCGGAGGTCTATGTACAGACCTT 242
QY 669 CAAGGCGCTCGGAGATTTCGGTGGACTACGTGCAGGTCTGT 708
|
|
|
Db 241 CAAGGCGCTCGGAGATTTCGGTGGACTACGTGCAGGTGAGT 202

RESULT 6
US-09-949-016-3728
; Sequence 3728, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3728
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3728

Query Match 25.5%; Score 272.8; DB 4; Length 1662;
Best Local Similarity 61.0%; Pred. No. 2.4e-45;
Matches 533; Conservative 0; Mismatches 302; Indels 39; Gaps 4;

QY 61 CGGAGGGTGTGACCATCGGCTTCGCGCTGTCTCATCTCGGCGCTCATGACCTGGGCGCTAC 120
|
|
|
Db 28 CGTGTGTGGGCACACGCTGTTTTCCTTGGCGTGTGGTGGCATCTTGGCAGCTAT 87
QY 121 GCGCGCGGGTGGCGCTGGCGCTCGGATCGCTACGGCTCTCGGCTTCGGCTCTACGG 180
|
|
|
Db 88 GTGACGGGCTACCAAGTTTCATCCACACGGAAGACACTACCTGTCTTCGGCTGTACGGC 147
QY 181 GCCTTCCTTTTCAGCGCACCTGTGGCGCAGAGCTCTTCGCTACCTGGAGCACCGCGG 240
|
|
|
Db 148 GCATCTCTGGGCGCTGACCTGTCTCATTCAGAGCTTTTTCCTTCCTGGAGCACCGGCGC 207
QY 241 GTGCGCGCGCGCGCGCGCGCGCGCTGGATGCAACCGCGC---GCAGTGTGGCGCTG 297
|
|
|
Db 208 ATGCGAGCTGCGCGCCAGGCGCTGAAGTGCCTCCCGCGCGCGGCTCGGTGGCACTG 267
QY 298 ACATCTCTCGCTACAGAGAGACCCCGCTACTGCGCAGTGCCTGGCGTCCGCGCGC 357
|
|
|
Db 268 TGCATTCCGCGCATACCAAGAGAGACCTCTGACTTTCGCAAGTGCCTTCGCGTCCGCGCAG 327
QY 358 GCGCTGTGTACCGCGCGCGCGCTGCGCTCTCATGTGTGGTGGATGGCAACCGCGCC 417
|
```

Db 328 CGCATCTCTCTCCCTGAC-----CTCAAGTGGTTCATGTTGGTGGTGAACGCCAG 381
Qy 418 GAGGACCTCTACATGGTGCACATGTTCCGCGAGGTCCTCGCTGACGAGGACCCGCCACG 477
Db 382 GAGGACGCTACATGCTGACATCTTCACGAGGTCCTGGCGGACCGAGGACCGCGC 441
Qy 478 TACGTGTGGGAGGCAATACACACAGCCCTGGGAAACCGCGCGCGCGCGCGTGGGC 537
Db 442 TTCTTTGTGTGGCGCAGCAACTTCCA-----TGAGGCGAGGCGAGGTGAG 486
Qy 538 GCCGAGGCTATCGGAGGTCGAGGCGGAGGATCTGGCGGCTGGCAGTGGAGCGCTG 597
Db 487 ACGGAGGCGCAGCTTCGAGAGGCGCATGACCGT-----GTGCGGATGTG 531
Qy 598 GTGAGGACTCGCAGGTGGCTGTGCTGGCGCAGCGCTGGGCGGCAAGCGGAGGTCTATG 657
Db 532 GTGCGGCGCAGCACCTTCTGTGTCATGTCGAGAGTGGGAGGCGAGGTCATG 591
Qy 658 TACACAGCTTCAAGCGCTCGGAGATTCGGTGGACTAGTGGAGTCTGTGACTCGGAC 717
Db 592 TACACGGCTTCAAGCGCTCGGCGATTCGGTGGACTACATCCAGTGTGGACTCTGAC 651
Qy 718 ACAAGTTGGACCCCATGCTGCTGGAGCTCGTGGGCTACTGGACGAGGACCCCGG 777
Db 652 ACTGTGTGGATCCAGCTCGACCATCGAGATGCTTCGAGTCTTCGAGGAGGATCCCAA 711
Qy 778 GTAGGGGCTGTGTGGGAGAGTGGCGATCTTAACCTCTGGACTCTGGGTTCAGCTTC 837
Db 712 GTAGGGGAGTGGGGGAGATGTCAGATCTCAACAGTAGTACGATCATGATTCCTTC 771
Qy 838 TAAAGCAGCTCGGATACCTGGGTAGCTTCAATGTGGAGCGGGCTGTGACAGTACTTC 897
Db 772 CTGAGCAGCTGGCGTACTGGATGGCTTCAAGTGGAGCGGCTGCCAGTCTACTTT 831
Qy 898 CACTGTGTATCTGTCATCAGCGTTCCTAGTA 931
Db 832 GGCTGTGTGCTGATGTAATTAGTGGGCCCTTGGGCA 865

RESULT 7

US-08-812-008-31
; Sequence 31, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustine, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwezman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665

; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1662
; OTHER INFORMATION:
US-08-812-008-31

Query Match 22.1%; Score 236.2; DB 4; Length 1665;
Best Local Similarity 58.4%; Pred. No. 4.5e-38;
Matches 512; Conservative 0; Mismatches 323; Indels 42; Gaps 4;

Qy 61 CGGAGGTCGTGACCATCGCTTCGCCCTTCCTCGCTCATCTCGGCTCATGACCTGGGCTTAC 120
Db 28 CGTGTGGTGGCACCAGTCTGTTTGGCTTGGTGGAGGATCCTTGGCGGCTAT 87
Qy 121 GCCCGCGGGTGGCGCTCGCATCGCTACGGCTTCCTGGCTTCCTGGCTTCGCGCTCTACGGG 180
Db 88 GTGACAGCTACCAAGTTTATCCACACAGAAAGACACTACTGCTCTTGGCTCTACGGT 147
Qy 181 GCCTTCCTTTTCAGCGCACTCGTGGCGCAGAGCTCTTCGCTACCTCGAGCACCGGCGG 240
Db 148 GCCATCTCTGGGTACATCTCTCATCCAGAGCTGTTTGGCTTCTCTGGAGCACGTCGA 207
Qy 241 GTGGCGGGCGGCGCGG-----GGGGCGCTGGATGACGACACCGCGCGGAGTGTGGG 294
Db 208 ATGCGCAGGGCAGGGCGGCGGCTCAAGCTGCATGCTCTCCAGAGTCCGCTTCAGTGGCA 267
Qy 295 CTGACCATCTCCGCTACAGAGAGACCCCGCTACCTCGCGCAGTGCCTGGCGTCCGCC 354
Db 268 CTCTGCATTTGCTTACCAAGAGACCCCGAATACCTGCGGAGTGCTTCGCTCAGCT 327
Qy 355 CGCGCCCTGTGTATCCCGCGCGCGGCTGCGCGCTCTCATGTGTGGTGGTGAACCGC 414
Db 328 CAGCGCATTTGCTTTCCAAAC-----CTCAAGGTGGTTCATGTAGTGGTGAATCGC 381
Qy 415 CGCGAGGACCTCTACATGGTTCGACATGTTCCGCGAGGTCCTTCGCTGACGAGGACCCGCC 474
Db 382 CAGGAAGATACCTACATGTTGGACATCTTCCATGAGGTGCTGGGTGGCACTGAGCAAGCT 441
Qy 475 ACGTACGTGTGGGAGCGCAACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGGTG 534
Db 442 GGCTTCTTGTGTGGCTAGCAATTTCCA-----TGAGCGGGGTGAAGA 486
Qy 535 GCGCCCGAGCCTATCGGAGGTCGAGCGGAGGATCCTTGGCGGCTGGCAGTGGAGCG 594
Db 487 GAGACAGAGGCGCCTGCGAGGAGGATGAGCGT-----GTGCGAGCT 531
Qy 595 CTGTGTAGGACTCGCAGGTGGTGTGGTGGCGCAGCGCTGGGGCGGCAAGCGGAGTGC 654
Db 532 GTGGTGTGGGCGAGCACCTTCTCATGTCATATGAGAAAGTGGGGGGGCAAGCGTGAAGTC 591
Qy 655 ATGTACACAGCTTCAAGGGCTCGGAGATTCGGTGGACTACGTGACGTCGTGACTCG 714
Db 592 ATGTACACTGCTTCAAGGCCCTTGGCAACTCAGTGGACTACATCCAGTGTGTGACTCT 651
Qy 715 GACACAAAGGTTGGACCCCATGCTGTGAGCTCGTGGGGTACTCGGAGGAGGACCCC 774
Db 652 GACACTGTGTGGACCCAGCCCTGCGACCATTCAGATGCTTCGAGTCTTGGAGAGATCCC 711
Qy 775 CGGTAGGGGCTGTGTGGGAGAGTGGCGGATCTTAAACCTCTGGAATCTCTGGGTGAGC 834
Db 712 CAAGTAGGAGGTGTGGAGGAGATGTCCAAATCCTCAACAAGTATGATTTCATGATCTCC 771

Qy 835 TTCCTAAGCAGCCTGGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTTCAGAGCTAC 894
Db 772 TTCTGTAGCAGTGTGAGTACTGGATGCTTTCAACGTGGAGCGGCGCTGCCAGTCTCTAC 831
Qy 895 TTCCACTGTGTATCTCTGCATCAGCGGTTCTCTAGGTA 931
Db 832 TTTGGCTGTGCAATGTATTAGTGGGCTTTGGCA 868

RESULT 8

US-09-949-016-201780/c
; Sequence 201780, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201780
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201780

Query Match 21.5%; Score 230.8; DB 4; Length 601;
Best Local Similarity 98.3%; Pred. No. 4.8e-37;
Matches 232; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 696 CGTCAGGCTGTGACTCGGCACAAAGTTGGACCCCATGTCAGTGTGGAGCTCGTGGCG 755
Db 403 CTTACAGGCTGTGACTCGGCACAAAGTTGGACCCCATGTCAGTGTGGAGCTCGTGGCG 344
Qy 756 GGTACTGACAGAGACCCCGGGTAGGGCTGTGTGGAGCTCGGATCCCTTAACCC 815
Db 343 GGTACTGACAGAGACCCCGGGTAGGGCTGTGTGGAGCTCGGATCCCTTAACCC 284
Qy 816 TCTGAGCTCTCTGGGTACGCTTCTAAGCAGCCTCGGATCTGGGTAGCCTTCAATGTGGA 875
Db 283 TCTGAGCTCTCTGGGTACGCTTCTAAGCAGCCTCGGATCTGGGTAGCCTTCAATGTGGA 224
Qy 876 GCGGGCTGTGAGAGCTACTTCCACTGTGTATCTGTATCTGATCAGCGGTTCTTAGGTA 931
Db 223 GCGGGCTGTGAGAGCTACTTCCACTGTGTATCTGTATCTGATCAGCGGTTCTTAGGTA 168

RESULT 9

US-08-675-499A-1
; Sequence 1, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schweigman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA

ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2947 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-675-499A-1

Query Match 16.5%; Score 176.2; DB 4; Length 2947;
Best Local Similarity 55.6%; Pred. No. 3.8e-26;
Matches 430; Conservative 0; Mismatches 238; Indels 45; Gaps 3;
Qy 159 CTTGGCCTTCGGCCTCTACCGGGCCTTCTTTCAGCGCAGCTGTGTGGCGAGAGCCTCTT 218
Db 636 CTTCTCATTTTGGACTGTACGCTGCTTTTGTAGCCTCGCATCTCATCATCCAAAGCCTCTT 695
Qy 219 CGGTACTCTGAGCACCGCGGGTGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCG 278
Db 696 TGGCTTTTGGACACCGGAAATGAAGAGTCCCTTGAACCCCGATTAATTTGAACAA 755
Qy 279 CGCGCGCAGTGTGGCGCTGACCATCTCCGCTTACAGAGAGACCCCGCTGACCTGCGCCA 338
Db 756 AAGC-----GTAGCACTCTGCATCGCTGCTGACCAAGAGAGACCTGACTACTTACGAA 809
Qy 339 GTGCTGGCTGCGCCCGCCCTGTGTGTACCGCGCGCGCGCTGCGCGCTCTCATGGT 398
Db 810 ATGTTTCAATCTGTGAAAGGCTGACCTACCCCTG-----GGATTAAAGTCGTGATGGT 863
Qy 399 GGTGGATGGCAACCGCGCGAGGACCTCTACATGGTTCGACATGTTCCGCGAGGCTTCGC 458
Db 864 CATGATGGGAACCTCAGACGACGACCTTTACATGATGACATATTCAGCGAAGTTATGG 923
Qy 459 TGACGAGGACCCCGCCACGCTACGTTGGGACGGCAACTACCACGAGCCTGGGAACCGCG 518
Db 924 CAGGACAAATCGGCCACGTACATCTGGAAGAACAACTTTCATGAAG----- 972
Qy 519 GGGCGGCGCGGTGGGCGCGGAGCTATTCGGGAGGTGGAGCGGAGGATCTCTGGCG 578
Db 973 -----GGACCTGGTGACACAGAGAGTCCCATGAAGAAAGTTC 1010
Qy 579 GTGGCAGTGGAGCGCTGTGTGAGGACTCGCAGGTGCTGTGCGTGGCGCAGCGCTGGGG 638
Db 1011 ACAACATGTACCCAAATGGTCTTGTCTAAACAAAGTATTTGCAATCATGCAAAATGGGG 1070
Qy 639 CGGCAAGCGGAGGTCACTATACACAGCCTTCAAGCGCCTCGGAGATTCGGTGGACTACGT 698
Db 1071 TGGAAAGAGAGAGTCACTATACACAGCCTTTCAGAGCACTGGGGCGAAGCGTGGATTATGT 1130
Qy 699 GCAGGTCTGTGACTCGGACACAGGTTGGAGCCCATCGCACTGCTGGAGCTCGTGGCGGT 758
Db 1131 ACAGGTGTGACTCAGATACATATGCTTGACCTTGCCTCATCTGTGGAGATGGTGAAGGT 1190

SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132199
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-132199

Query Match 12.8%; Score 137.4; DB 4; Length 601;
Best Local Similarity 58.5%; Pred. No. 1.6e-18;
Matches 300; Conservative 0; Mismatches 201; Indels 12; Gaps 3;

```
QY 61 CGGAGGTGCTGACCATCGCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
DB 50 CGTGTGTGGGACCAAGCAGCTGTGTTGCCCTGGCAGTGTCTGGTGGCATCTCTGACGCTAT 109
QY 121 GCCGCGGGGTGCGCTGCGCTCCGATCGCTAGCGCTCTCTGGCTTCTGGGCTCTACGGG 180
DB 110 GTGACGGGTACCACTTATCATCAGGAAAGACTACTCTCTCTGGGCTCTACGGC 169
QY 181 GCCTTCTTTTTCAGCGACCTGCTGTGGCGAGAGCTCTTTTGGCTACTTGGAGCACCGGG 240
DB 170 GCATCTCTGGGCTGCACTGCTCATTCAGAGCCCTTTTGGCTTCTCTGGAGCACCGGCG 229
QY 241 GTGGCGGGCGGGCGGGGCGCTGTGATGAGCACCGGCGC----GCAGTGTGGCGCTG 297
DB 230 ATGCGACGTGCGGGCGAGCCCTGAAGCTGCCCTCCCGCGGGGCTCGGTGGCACTG 289
QY 298 ACCATCTCCGCTACACAGAGAGACCCGCTACCTCGCGCAGTGCCTGCGGTCCGCCCGC 357
DB 290 TGATTTGCGGCTTACAGAGAGACCTGACTACTTGGCAAGTGCCTGCGCTCGGCGCCAG 349
QY 358 GCCCTGCTGTACCCGCGCGCGGCTGCGGCTGCTCATGCTGTGGTGGATGGCAACCGCGCC 417
DB 350 CGCATCTCTTCCCTGAC-----CTCAAGTGTGTCATGTTGATGGTGGCAACCGCCAG 403
QY 418 GAGGACCTTACATGCTGACATGTTTCCGCGAGGCTT---CGCTGACGAGAGACCCCGCC 474
DB 404 GAGGACGCTTACATGCTGACATCTTCCACGAGTGTCTGGGCGGACACGAGCAGCGCGC 463
QY 475 ACGTACGTGTGGAGCGCACTTACCACGACCCCTGGGAACCGCGCGGGCGGCGCGGTG 534
DB 464 TTTCTTGTGTGGCGGACGACCTTCCATGAGGAGCGAGGCTGAGACGAGCCAGCCTG 523
QY 535 GCGCGCGGAGCCTATCGGAGGTGAGGCGGAG 567
DB 524 CAGGAGGGCATGGACCGTGTGGGATGTGGT 556
```

RESULT 14

US-09-902-540-8819
; Sequence 8819, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8819
; LENGTH: 6645
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-8819

Query Match 7.7%; Score 82.8; DB 4; Length 6645;
Best Local Similarity 44.8%; Pred. No. 1.4e-07;
Matches 357; Conservative 0; Mismatches 437; Indels 2; Gaps 1;

```
QY 3 GAGACAGCAGCAGCGCGCCCAAGCCACATCTCTGACAGCCGCGCTCTCTCGGCTTGGCCG 62
DB 3823 GAGGACGCGCCACCGCGCGAGGTGCTTACCGCGCGCGCGGAGCTCTTTCGACGACCG 3882
QY 63 GAGGTGTGCTGACCATTCGCGC--TTGCGCTTGTCTATCTTGGGCTCATGACCTTGGGCTTAC 120
DB 3883 AAGGACGCAATCGAGCGCTGTGTCGCTGGCGCGCTGCGCTCCCGCGAGATGCTGCTCCGTC 3942
QY 121 GCCGCGGGGTGCGCTGCGCTCCGATCGCTAGCGCTTCTGGCTTCTGGGCTTCTGGGCTTACGG 180
DB 3943 ATCGACCGGGCGGTGCGCGGGGTGCACTCTGGGAGCTCATGGGAGCTCTGCTTGGACGTG 4002
QY 181 GCCTTCTTTTTCAGCGCACCTTGTGGCGCAGAGCTCTTTCGGGTACTTGGAGCACCGCGGG 240
DB 4003 TAGCGCTCGGCGCGGAGCGCGGTGCGCTCGCGCGCGGAGCTCTGTTTGGCGCGC 4062
QY 241 GTGGCGGGCGCGCGGGGCGCTGTGATGAGGACCGCGCGGAGTGTGGCGCTGACCGTAC 300
DB 4063 CGCTGCTGGGCTTCCAGCTCGCTGGCGGATGAAGACGCGGCTTGGAGCTTGAACCGAGCGC 4122
QY 301 ATCTCCGCTTACAGGAGGACCCCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCGCGCGCGC 360
DB 4123 GCCGCGGAGCGCGCGCGGAGGACCTTACCGGCTTCAAGGACTGTGTGACGGGCTTGGGT 4182
QY 361 CTGCTGTACCGCGCGCGCTGCGCTCTCTCATGTGTGTGGATGGCAACCGCGCGCGAG 420
DB 4183 CAGCGGGCGACTCCACGCGGCTGTGGAGCGCTGGAGCAGTGGGTGCGCGCGCTCGAG 4242
QY 421 GACTCTTACATGTGTGACATGTTCCGCGAGTGTTCGTGTGACGAGGACCCCGCGCTACGTAC 480
DB 4243 GACGCGGACGAGGCGCGGCTGCTGCGCTGAGGCTTGGCGGAGCTGCGCGCTGACCGCGCGC 4302
QY 481 GTGTGGAGCGCACTACACAGCCCTGGGAACCGCGCGGGCGGCGCGGTGGGCGCGC 540
DB 4303 CGGTGTGATGTGGCGCGGAGGACTGTGAGGAAAGTGGCGCGGCTTGGGCGCTTGGGCGCG 4362
QY 541 CGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTTGGCAGTGGAGCGCTGTGTG 600
DB 4363 GGTACGCGGAGCGCTGGAAGCCTCGAGCGCTGCTGAGAGCGCGCGCGCGCGCGCGC 4422
QY 601 AGGACTCGCAGGTGTGTGCGCGCAGCGCTGCGGCGGCGGAGCGGAGGCTATGTAC 660
DB 4423 CCGGAGGTGAGTGTGGCGCGGAGTGTGCGCTTCCGCGCGGAGCGGAGGCTGTGCTC 4482
QY 661 ACAGCTTCAAGGCGCTCGGAGATTTCGTGACTACGTGAGGCTGTGTGACTCGGACACA 720
DB 4483 ATGGCGCGCGCGGGGCTTTCAGTTCGCGGAGACAGCTTTCGAGGCGGCTTGAAGCGCGCG 4542
QY 721 AGGTGGACCCCATGGCACTCTGTGAGCTCTGTGGGTACTTGGAGGAGACCGCGGGGTA 780
DB 4543 AAGGACGCGCTCGGTGGAGCGGAGCGTTCGAGCGCGCTTGGAGTGGCGCGCGCTGTAC 4602
QY 781 GGGGCTGTGTGGG 796
DB 4603 CGCGCTCGGCGGAGG 4618
```

RESULT 15

US-09-902-540-938
; Sequence 938, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 938
; LENGTH: 10210
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-938

```

Query Match      7.7%; Score 82.8; DB 4; Length 10210;
Best Local Similarity 44.8%; Pred. No. 1.5e-07;
Matches 357; Conservative 0; Mismatches 437; Indels 2; Gaps 1;

QY 3 GAGACAGCAGGACGCGCCAAAGCCACTCCTGACGCGCGCTGCTCGGCGCTGGCCCG 62
Db 4198 GAGGACGCGCCACCGCGAGGTGCTTACCGCGCGCGGGACCTTTTCGACGACCG 4257

QY 63 GAGGTGCTGACCAATCGCC--TTCCGCTCTGCTCATCTTGGGCTCATGACCTGGGCTTAC 120
Db 4258 AAGGACGCGCATCGACGCGCTGTTCCGCTGGCGGCTTGGTCCCGGATGCGTCCGTC 4317

QY 121 GCGCGCGGGGTGCGGCTCGCTCCGATCGCTACGGCTCTTGGGCTTTCGGGCTTACGGG 180
Db 4318 ATCGACCGGCGGCTGGCGGGGTGCACGCCCTTGGGACGTCATGGCGACCTGCTGGACGTG 4377

QY 181 GCCTTCCTTTTCAGCGCACTGCTGGTGGCGAGAGCCCTTTCGGGTACTCTGAGCACCGGCGG 240
Db 4378 TACGCTCGGGCGCGAGCGCGGGTGGCGCTCGCGCGCGGAGCTGCTGTGGCCGCC 4437

QY 241 GTGGCGGCGGCGCGCGGGCGCTGGATGAGCACCGCGCGCAGTGTGGCGCTGACC 300
Db 4438 GCGTCGGTGGCTCCAGCTCGCTGGCGGATGAAGACGCGGCTGGACGCTGACCCAGCGC 4497

QY 301 ATCTCCGCTTACAGAGGACCCCGGTAACCTGGCGGAGTGGCTGGCGCTCGCCCGCGCC 360
Db 4498 GCCGCGGAGCGCGCGCGGAGACCTCACCGGCTTCCAGGACCTGGTGTACCGGGCTGCGT 4557

QY 361 CTGCTGTATACCGCGCGCGCTGGCGCTCTCTCATGTGGTGGATGGCAACCGCGCGCGAG 420
Db 4558 CAGCGGGCGGACTCACCGGCTGCTGAGGGGCTGGAGCAGTGGGTGCCGGGCTCGAG 4617

QY 421 GACCTCTATATGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC 480
Db 4618 GACGCGGACGAGCGCGCGCTGCGGCTGGAGCTGGCGGAGCTCGCGCTGACGCGGGC 4677

QY 481 GTGTGGGACGGCACTACACAGCCCTGGGNAACCGCGGCGGGCGCGGCTGGGCGCC 540
Db 4678 CCGTGGATGTGGCGCGCGAGGCACTGGAGGAAGTGGCGCGCGCTCGGGCGCG 4737

QY 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCCTGGGCGGCTGGCAGTGGAGGCGCTGGTG 600
Db 4738 GGTTACGGGACGCGCTGGAGCCCTCGAGCGCTGCTGAAGAGCGCCCCCGCGCGCGC 4797

QY 601 AGGACTCGAGTGTGCTGGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTGATGTAC 660
Db 4798 GCCGAGGTGCAAGTGGCGCGCGAGTTGGCTTCCGCGCGGAGCGCAGGTGCTGCTC 4857

QY 661 ACAGCCTTCAAGGCGCTCGAGATTCGGTGAACACTGCTGAGGTCTGTGACTCGGACACA 720
Db 4858 ATGGCGCGCGCGGGGCTTCGAGTTCGCGGAGACAGCTTCCGGAGGGCCCTGAAAGCGCGG 4917

QY 721 AGGTTGGAACCCATCGGCTGCTGGAGCTCTGTGGGGTACTGGACGAGGACCCCGGGTA 780
Db 4918 AAGGACGCGGCTCGGTGGAGCGGACGCTCGACGCGGCGCTTGGAGTGGCCACCTGTAC 4977

QY 781 GGGGCTGTGTGGG 796
Db 4978 CGCGCTCGGCGAGG 4993

```

Search completed: March 13, 2005, 08:12:15
Job time : 217.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:08:52 ; Search time 4414.94 Seconds
(without alignments)
9233.835 Million cell updates/sec

Title: US-10-672-399-3
Perfect score: 1071
Sequence: 1 atgagacagcaggacgcgc.....ccggcaccatcgctggatga 1071

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	916.8	85.6		2072	3	CR602106	
2	915.2	85.5		1728	9	AY421569	Homologous
3	822.4	76.8		1058	4	BM543754	Agonotetrax
4	795.6	74.3		965	1	AL545447	
5	679.8	63.5		1746	9	AY421571	Mus musculus
C 6	490.4	45.8		896	5	BX371636	
7	479.4	44.8		3542	3	AK053726	Mus musculus
8	357.4	33.4		650	2	BM665418	
9	305.8	28.6		1084	4	BM544718	Agonotetrax
10	291.8	27.2		551	4	EG707429	
11	291.8	27.2		551	4	EG707608	
12	268.2	25.0		508	9	CE377866	tigr-g88
13	258.2	24.2		610	5	BX327795	BX327795
14	256.2	23.9		894	7	CF239758	Agonotetrax
15	255.8	23.9		782	4	BI753116	
C 16	253.6	23.7		668	8	BM057613	
17	237.8	22.2		774	5	BP677403	BP677403
18	236.2	22.1		4151	3	AK028582	Mus musculus
19	223.8	20.9		895	5	BU911059	Agonotetrax
20	209	19.5		747	6	CD466132	Leukon2_2
21	206.8	19.3		604	4	EG016146	Agonotetrax
22	202.6	18.9		944	6	CA487249	
C 23	194.2	18.1		495	7	CN664556	
24	186.2	17.4		660	5	BP674137	BP674137

25	186.2	17.4	918	7	CF220501
26	182.4	17.0	739	9	AY421570
27	179.6	16.8	739	7	CN304087
28	176.2	16.5	2891	3	AK079729
29	176	16.4	653	7	CV029395
30	173.2	16.2	857	5	BU906272
31	173.2	16.2	866	6	CA988122
32	172.4	16.1	848	6	CA987283
33	171	16.0	809	5	BP693553
34	170.6	15.9	748	5	BP699960
35	165.4	15.4	781	5	BP693535
36	164.8	15.4	746	5	BP705367
37	164.6	15.4	871	5	BU910414
38	163.6	15.3	920	7	CF220396
39	162.8	15.2	840	5	BU911086
40	162.8	15.2	941	6	CA790329
41	162.6	15.2	647	7	CN666988
42	162.6	15.2	853	6	CA988392
43	162.6	15.2	864	5	BU907701
44	162.6	15.2	898	5	BU906701
45	162.4	15.2	923	5	BU909569

ALIGNMENTS

RESULT 1
CR602106
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
REFERENCE
AUTHORS
JOURNAL
COMMENT
FEATURES
ORIGIN
Query Match
Best Local Similarity
Matches
QY
DB

2072 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CSODI027YJ02 of Placenta Cot 25-normalized
of Homo sapiens (human).
CR602106
CR602106.1 GI:50482913
HTC; CDSLT cDNA.
Homo sapiens (human)
Homo sapiens
Homo sapiens
1 (bases 1 to 2072)
Li W.B., Gruber C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2072)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequenage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1..2072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI027YJ02"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

85.6%; Score 916.8; DB 3; Length 2072;
99.8%; Pred. No. 2.6e-174;
918; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
10 CAGGACGGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCGGCTCGGCGGAGGTG 69
58 CAGGACGGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCGGCTCGGCGGAGGTG 117

```
Qy 70 CTGACCATCGCTTCGGCCTCTCATCTCGGCTCATGACTGGGCTATCGCCGCGG 129
Db 118 CTGACCATCGCTTCGGCCTCTCATCTCGGCTCATGACTGGGCTATCGCCGCGG 177
Qy 130 GTGCCGCTGGCTTCGATCGCTACGGCTCTCGGCTTCGGCTCTACGGGCGCTTCCTT 189
Db 178 GTGCCGCTGGCTTCGATCGCTACGGCTCTCGGCTTCGGCTCTACGGGCGCTTCCTT 237
Qy 190 TCAGCGCACTGTGTGGCGAGAGCTCTTCGCTACCTTGGAGCAACCGCGGCTGGCGCG 249
Db 238 TCAGCGCACTGTGTGGCGAGAGCTCTTCGCTACCTTGGAGCAACCGCGGCTGGCGCG 297
Qy 250 GCGCGCGCGGCGCTGTGATGACGACCGCGCGCGAGTGTGGCTGACCATCTCGCC 309
Db 298 GCGCGCGCGGCGCTGTGATGACGACCGCGCGCGAGTGTGGCTGACCATCTCGCC 357
Qy 310 TACAGAGAGACCGCGCTACTCGCGCAGTGCCTGGCGTCCGCGCGGCTCTGTGTAC 369
Db 358 TACAGAGAGACCGCGCTACTCGCGCAGTGCCTGGCGTCCGCGCGGCTCTGTGTAC 417
Qy 370 CCGCGCGCGGCTGCGCTCTCATGTGTGTGATGCAACCGCGCGAGGACTCTTAC 429
Db 418 CCGCGCGCGGCTGCGCTCTCATGTGTGTGATGCAACCGCGCGAGGACTCTTAC 477
Qy 430 ATGCTCGACATGTTCCGCGAGGTCTTCTGCTGACGAGGACCGCGCAGTACGTGTGGAC 489
Db 478 ATGCTCGACATGTTCCGCGAGGTCTTCTGCTGACGAGGACCGCGCAGTACGTGTGGAC 537
Qy 490 GGCAACTACCAACAGCCCTCGGAAACCGCGCGCGCGGCTGGGCGCGGAGCCTAT 549
Db 538 GGCAACTACCAACAGCCCTCGGAAACCGCGCGCGCGGCTGGGCGCGGAGCCTAT 597
Qy 550 CGGAGGTGTGAGGCGGAGGATCTTGGCGGCTGTGAGTGGAGCAACCGCGCGAGTGTG 609
Db 598 CGGAGGTGTGAGGCGGAGGATCTTGGCGGCTGTGAGTGGAGCAACCGCGCGAGTGTG 657
Qy 610 AGGTGCGTGTGCGTGGCGCAGGCTTGGCGGCGGCGGCGGCTGGGCGCGGAGCCTTC 669
Db 658 AGGTGCGTGTGCGTGGCGCAGGCTTGGCGGCGGCGGCGGCTGGGCGCGGAGCCTTC 717
Qy 670 AAGCGCTTCGAGATTCGGTGGACTTACGTGAGGTCTGTGACTCGGACACAAAGGTGGAC 729
Db 718 AAGCGCTTCGAGATTCGGTGGACTTACGTGAGGTCTGTGACTCGGACACAAAGGTGGAC 777
Qy 730 CCCATGACATGCTGTGAGCTCGTGGGTAATGACGAGGACCCCGGGTATGGGCTGTT 789
Db 778 CCCATGACATGCTGTGAGCTCGTGGGTAATGACGAGGACCCCGGGTATGGGCTGTT 837
Qy 790 GGTGGGACGTGGGATCCTTAACCTCTGACCTTCTGAGTCTTGGGTGAGTCTTAAGCAGCTG 849
Db 838 GGTGGGACGTGGGATCCTTAACCTCTGAGTCTTGGGTGAGTCTTAAAGCAGCTG 897
Qy 850 CGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCACTGTGTATCC 909
Db 898 CGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCACTGTGTATCC 957
Qy 910 TGCATCAGCGGTTCTCTAGG 929
Db 958 TGCATCAGCGGTTCTCTAGG 977
```

RESULT 2

```
AY421569
LOCUS AY421569 1728 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY421569
VERSION AY421569.1 GI:39748428
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 1728)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302

2 (bases 1 to 1728)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES

source
1..1728
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene
1..1728
/gene="HAS1"
/locus_tag="HOM7601"

ORIGIN

Query Match 85.5%; Score 915.2; DB 9; Length 1728;
Best Local Similarity 99.7%; Pred. No. 5.4e-174;
Matches 917; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy 10 CAGGACCGCGCCAAAGCCCACTCTCTGACGCGCGCGCTGCTCGGCTTACGGGCGGAGGTG 69
Db 1 CAGGACCGCGCCAAAGCCCACTCTCTGACGCGCGCGCTGCTCGGCTTACGGGCGGAGGTG 60
Qy 70 CTGACCATCGCTTCGGCTCTCATCTGGGCTCATGACTGGGCTTACGGGCGGCGG 129
Db 61 CTGACCATCGCTTCGGCTCTCATCTGGGCTCATGACTGGGCTTACGGGCGGCGG 120
Qy 130 GTGCGCTGCGCTTCGGCTCTCATCTGGGCTCTCGGCTTACGGGCGGCTTCTT 189
Db 121 GTGCGCTGCGCTTCGGCTCTCATCTGGGCTCTCGGCTTACGGGCGGCTTCTT 180
Qy 190 TCAGCGCACTGTGTGGCGAGAGCTCTTCTGCGTACCTTGGAGCACCGCGGCGGCGG 249
Db 181 TCAGCGCACTGTGTGGCGAGAGCTCTTCTGCGTACCTTGGAGCACCGCGGCGGCGG 240
Qy 250 GCGCGCGGCGGCGCTGTGATGACGACCGCGGCGAGTGTGGGCTTACCATCTCGCC 309
Db 241 GCGCGCGGCGGCGCTGTGATGACGACCGCGGCGAGTGTGGGCTTACCATCTCGCC 300
Qy 310 TACCAGAGGACCCCGGCTACCTGCGCAGTGTGCTGCGCTCCGCGCGGCTCTGTAC 369
Db 301 TACCAGAGGACCCCGGCTACCTGCGCAGTGTGCTGCGCTCCGCGCGGCTCTGTAC 360
Qy 370 CCGCGCGGCGGCTGCGGCTCTCATGTGTGTGATGGCAACCGCGCGGAGGAGCTCTAC 429
Db 361 CCGCGCGGCGGCTGCGGCTCTCATGTGTGTGATGGCAACCGCGCGGAGGAGCTCTAC 420
Qy 430 ATGCTCGACATGTTCCGCGAGGTCTTCTGCTGACGAGGACCCCGGCACTGCTGTGGAC 489
Db 421 ATGCTCGACATGTTCCGCGAGGTCTTCTGCTGACGAGGACCCCGGCACTGCTGTGGAC 480
Qy 490 GGCAACTTACCACCAAGCCCTCGGAAACCGCGCGGCGGCTGGGCGGCGGAGCCTAT 549
Db 481 GGCAACTTACCACCAAGCCCTCGGAAACCGCGCGGCGGCTGGGCGGCGGAGCCTAT 540
Qy 550 CCGGAGGTGTGAGGCGGAGGATCTTGGGCGGCTGTGAGGCGGCTGTGTGAGACTCGC 609
Db 541 CCGGAGGTGTGAGGCGGAGGATCTTGGGCGGCTGTGAGGCGGCTGTGTGAGACTCGC 600
```



```
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31267282.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?r=CS0DI027DE01QPI&c=9502.r.
FEATURES
    source
        1..965
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DI027JU02"
            /tissue_type="PLACENTA COT 25-NORMALIZED"
            /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized."
ORIGIN
    Query Match 74.3%; Score 795.6; DB 1; Length 965;
    Best Local Similarity 97.4%; Pred. No. 6.4e-150;
    Matches 851; Conservative 9; Mismatches 8; Indels 6; Gaps 5;
Qy 10 CAGAGCGCCCAAGCCACCTCTGCGAGCCGCGCTGCTCGGCTGCGCGGAGGGTG 69
Db 58 CAGAGCGCGCCCAAGCCACCTCTGCGAGCCGCGCTGCTCGGCTGCGCGGAGGGTG 116
Qy 70 CTGACCATCGCTTTCGCGCTCATCTGCGGCTCATGACCTGCGCTACGCGCGCGG 129
Db 117 CTGACCATCGCTTTCGCGCTCATCTGCGGCTCATGACCTGCGCTACGCGCGCGG 176
Qy 130 GTGCGGCTGCGCTTCGATCGCTAGCGGCTCTGCGCTTCGCGCTTCAGCGGCTTCCTT 189
Db 177 GTGCGGCTGCGCTTCGATCGCTAGCGGCTCTGCGCTTCGCGCTTCAGCGGCTTCCTT 236
Qy 190 TCAGCGCACTGTGTGGCGCAGAGCTCTTTCGCTACCTGTGAGCACCAGCGCGGTGGCGCG 249
Db 237 TCAGCGCACTGTGTGGCGCAGAGCTCTTTCGCTACCTGTGAGCACCAGCGCGGTGGCGCG 296
Qy 250 GCGCGCGCGCGCGCTGGATGAGCACCACCGCGCAGTGTGGCTGACCATTCGCGC 309
Db 297 GCGCGCGCGCGCGCTGGATGAGCACCACCGCGCAGTGTGGCTGACCATTCGCGC 356
Qy 310 TACCAGAGGACCCGCGCTACTCGCGCAGTGTGGCTGCGCGCGCGCTGCTGTAC 369
Db 357 TACCAGAGGACCCGCGCTACTCGCGCAGTGTGGCTGCGCGCGCGCTGCTGTAC 416
Qy 370 CCGCGCGCGCGCTGCGCGTCTCATGTGTGTGATGCGCAACCGCGCGGAGCACTCTAC 429
Db 417 CCGCGCGCGCGCTGCGCGTCTCATGTGTGTGATGCGCAACCGCGCGGAGCACTCTAC 476
Qy 430 ATGTGTCACATGTTCCGAGAGTTCCTGCTACCGAGNACCCGCGCAGTCTGTGGAC 489
Db 477 ATGTGTCACATGTTCCGAGAGTTCCTGCTACCGAGNACCCGCGCAGTCTGTGGAC 536
Qy 490 GGCAACTACCAACAGCCCTGGGAACCCGCGCGCGCGCGGTGGCGCGCGAGCCTAT 549
```

```
537 GGCAACTACCAACAGCCCTGGGAACCCGCGCGCGCGGTGGCGCGCGAGCCTAT 596
550 CCGAGAGTGGAGGCGGAGGATCCTGGCGGCTGCGAGTGGAGGCGCTGCTGAGGACTGCG 609
597 CCGAGAGTGGAGGCGGAGGATCCTGGCGGCTGCGAGTGGAGGCGCTGCTGAGGACTGCG 656
610 AGGTGCGCTGTGCGTGGCGCAGCGCTGGCGGCGCAAGCGCGAGGTGCTGTCACACAGCCCTTC 669
657 AGGTGCGCTGTGCGTGGCGCAGCGCTGGCGGCGCAAGCGCGAGGTGCTGTCACACAGCCCTTC 716
670 AAGCGCTCGGAGATTCGGTGGACTACGTGCGAGTCTGTGACTCGGACACAAAGGTGGAC 729
717 AAGCGCTCGGAGATTCGGTGGACTACGTGCGAGTCTGTGACTCGGACACAAAGGTGGAC 776
730 CCATGCGACTGCTGAGCTGCTGCGGCTACTGACAGGACCCCGGTTAGGCGCTTT 789
777 CCATGCGACTGCTGAGCTGCTGCGGCTACTGACAGGACCCCGGTTAGGCGCTTT 836
790 GGTGGGACGTGGCGATCCTTAACCTCTGAGCTCCTGGGTGAGTCTTCTAAGCAGCCTG 849
837 GGTGGGACGTGGCGAT-CTTAMCCTTGGACTCTTGGTAR--CTTCTAAGCAG-CTG 892
850 CGATACTGGGTAG-CTTAAATGTGAGCGGCTT 883
893 CGATACTGGGTAG-CTTAAATGTGAGCGGCTT 925

RESULT 5
AY421571 1746 bp DNA linear GSS 17-DEC-2003
LOCUS Mus musculus HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY421571
VERSION AY421571.1 GI:39748430
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1746)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1746)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
    source
        1..1746
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            <1..>1746
            /gene="HAS1"
            /locus_tag="HCM7601"
    gene
    ORIGIN
        Query Match 63.5%; Score 679.8; DB 9; Length 1746;
        Best Local Similarity 84.6%; Pred. No. 1.3e-126;
        Matches 794; Conservative 0; Mismatches 127; Indels 18; Gaps 2;
Qy 10 CAGGACGCGCCCAAGCCACTCTCTGACGCGCGCTGCTCGGCGCTGCGCGGAGGTG 69
```

```
Db 1 CAGGACATGCCAAGCCCTCAGAGCAGCGCTTCTGCTCTGGCCCTGCGCCAGCGGGCA 60
Qy 70 CTGACCATCGCTTCCGCTCTGCTATCTCTGGGCTCATGACCTGGGCTCATGCGCCGGG 129
Db 61 CTCAGATCATCTTGGCTCTGCTATCTCTGGGCTCATGACCTGGGCTCATGCGCCAGGC 120
Qy 130 GTGCGCTGGCTTCGATCGCTACGCGCTCTTGGCTCTTGGGCTCTTACGCGGCTTCTT 189
Db 121 GTTCTCTGGCTTCAGATCGCTATGAGCTCTTGGCTCTTATGGGCTCTATGGGCTTCTC 180
Qy 190 TCAGGACCTTGGTGGCGCAGAGCTCTTCCGCTACCTGGACACCGCGGGTGGCGGG 249
Db 181 AGCGACACCTAGTGGGACAGAGCTCTTCCGCTTACCTTGGACACCGAAGGCTGGCAGC 240
Qy 250 GCGGCGCG-----GGGCGCGTGTGATGACGACCGCGCGGAGTGTGGCG 294
Db 241 GCTGGCGCGCGCTCTTGGCGNAGGGGCGCTTGGATGGGCGCTGACGACGAGGTGGCA 300
Qy 295 CTGACCATCTCCGCTACAGAGAGACCCCGCTACCTCGCGCACTGCTGGCGTCCGCC 354
Db 301 CTCACCATCTCAGCTTACCAAGAGATCCCGCTTACCTCGCGCACTTGAACCTCCGCG 360
Qy 355 CGCGCTCTGTATCCCGCGCGCGCTGGCGCTCTTCCGCTCTCATGCTGGTGGATGGCAACGC 414
Db 361 CGCGCTTGTGTATCCCGCACACGAGGTTACGCGTCTCATGCTGGTGGTGGAGCGCAACGC 420
Qy 415 GCCGAGGACCTTACATGCTGCACATGTTCCGCGAGGCTTTCGCTGACGAGGACCCCGCC 474
Db 421 GCTGAGGATCTGTATGCTGACATGTTCCGAGAGTCTTCCGCGATGAGACCCCGCC 480
Qy 475 ACGTACGTGTGGGACGCACTACACAGAGCTTGGAAACCGCGCGCG---CGGCGCG 531
Db 481 ACTTATGTTGGATGGCACTTACCATCAGCCCTGGGAACACGAGCGGAGCTACGCGCGCT 540
Qy 532 GTGGCGCGGAGCTTATCGGAGGTGAGCGGAGGATCTTGGCGGCTGCGAGTGGAG 591
Db 541 GTCGCTGAAGGTGCTTACCGGAGGTGAGCGGAGGACCCCGCGCGTGGCGTGGAG 600
Qy 592 GCGCTGTGTAGGACTCGAGGTGCGTGTGGCGAGCGCTGGGGCGGCAAGCGCGAG 651
Db 601 GCGCTGTGTAGAACACAGCAGGTGCGTGTGGCTGAGCGTGGGGCGGCAAGCGTGGAG 660
Qy 652 GTCATGTACACAGCTTCAAGCGCTCGGAGATTCGCTGAGACTAGTGCAGGTCTGTGAC 711
Db 661 GTCATGTACACAGCTTCAAGGCACTGGCGACTCCGTGGACTAGTGCAGGTCTGTGAC 720
Qy 712 TCGGACACAGGTTGGACCCCATGCGACTGCTGAGCTCGTGGCGGTACTGACGAGGAC 771
Db 721 TCAGACACAGACTAGACCCCATGCGACTGCTGAGCTTGTGCGAGTGTGGATGAAGAC 780
Qy 772 CCCGGGTAGGGCTGTGGTGGGACGCTGGGATCTTAACTTAACTTAACTTAACTTAACTT 831
Db 781 CCCGGGTAGGGCTGTGGAGGGATGTGAGGATCTTAACTTAACTTAACTTAACTTAACTT 840
Qy 832 AGCTTCTTAAGCAGCTCTCGATPACTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGC 891
Db 841 AGCTTCTTGAAGCTTCTCGATPACTGGGTAGCTTCAATGTGGAGCGAGCTTGTGAGAGC 900
Qy 892 TACTTCCACTGTGTATCTCGATPACTAGCGGTCTCTAGGT 930
Db 901 TACTTCCACTGTGTCTCGATPACTAGCGGTCTCTAGGT 939
```

RESULT 6
BX371636/c
LOCUS BX371636 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI027YJ02 3-PRIME, mRNA sequence.
ACCESSION BX371636
VERSION BX371636.2 GI:46616231
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 896)
TITLE Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT On May 8, 2003 this sequence version replaced gi:30450029.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS0BA1019ZC12_CS01819_1&c=9502.r

FEATURES
source
Location/Qualifiers
1..896
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI027YJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
primer="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 45.8%; Score 490.4; DB 5; Length 896;
Best Local Similarity 92.8%; Pred. No. 1.7e-88;
Matches 553; Conservative 0; Mismatches 39; Indels 4; Gaps 4;
Qy 335 GCCAGTGCTCGCTCGCGCCGCTGCTGCTACCGCGCGCGCGCTGCGCGCTGCTCA 394
Db 896 GCCAGTGCTCGCTCGCGCCGCTGCTGCTGCTACCGCGCGCGCGCTGCGCGCTGCTCA 837
Qy 395 TGGTGTGGATGGCAACCGCGCCGAGGACCTCTACATGCTGCGACATGTTCCGCGAGGTCT 454
Db 836 TGGT-GTGGAGGACACCGCGCCGAGGACTTTACAATGCTGCGACA-GTTCCCGAGGTCT 779
Qy 455 TCGCTGACGAGGACCCCGCCACGTAAGTGTGGAGCGGCAACTACACAGCGCTGGG-AA 513
Db 778 TCGCTGACGAGGACCCCGCCACGTAAGTGTG-GACGCGCAACTACACAGCGCTGGGAAA 720
Qy 514 CCCGCGCGCGCGCTGCGCGCGCGCTGCGAGGCTTATCGGAGGTGGAGCGGAGGATCCT 573
Db 719 CCCGCGCGCGCGCTGCGCGCGCGCTGCGAGGCTTATCGGAGGTGGAGCGGAGGATCCT 660
Qy 574 GGGCGGCTGGCAGTGGAGCGCGCTGCTGAGGACTCGCAGGTGCTGCTGCGCGAGCGC 633
Db 659 GCGCGCTGGCAGTGGAGCGCGCTGCTGAGGACTCGCAGGTGCTGCTGCGCGAGCGC 600
Qy 634 TGGGCGCGCAAGCGCGAGGTCATGTACACAGCTTCAAGGCGCTCGAGAGATTCGCTGGAC 693
Db 599 TGGGCGCGCAAGCGCGAGGTCATGTACACAGCTTCAAGGCGCTCGAGAGATTCGCTGGAC 540
Qy 694 TACGTGAGGTCTGTGACTCGGACACAGGTTGACCCCATGGCACTCTGGAGCTCGTG 753
Db 539 TACGTGAGGTCTGTGACTCGGACACAGGTTGACCCCATGGCACTCTGGAGCTCGTG 480
Qy 754 CGGCTACTGGACGAGGACCCCGCGTGGGCTGTGTGGTGGGAGCTGCGGATCTCTTAAC 813
Db 479 CGGCTACTGGACGAGGACCCCGCGTGGGCTGTGTGGTGGGAGCTGCGGATCTCTTAAC 420
Qy 814 CCTCTGAGCTCCTGGGTGAGCTTCTTAAGCAGCTTCGATATCTGGGTAGCTTCAATGTG 873

Db 419 CCTCGACTCCTGGGTGAGCTTCTTAAGCAGCCTCGGATCGGATAGCTTCAATGTG 360

Qy 874 GAGCGGCTTCTGAGAGTACTTCCACCTGTATCTTCGTCATCAGCGGTCTCTAGG 929
|||||

Db 359 GAGCGGCTTCTGAGAGTACTTCCACCTGTATCTTCGTCATCAGCGGTCTCTAGG 304
|||||

RESULT 7

AK053726

LOCUS AK053726 3542 bp mRNA linear HTC 03-APR-2004

DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone: E130302P16 product: hyaluronan synthase1, full insert sequence.

ACCESSION AK053726

VERSION AK053726.1 GI:26343670

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 Carninci, P. and Hayashizaki, Y.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE 3
20493374
11042159

REFERENCE 4
20530913
11076861

TITLE The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

JOURNAL Functional annotation of a full-length mouse cDNA collection

MEDLINE Nature 409, 685-690 (2001)

PUBMED

REFERENCE 5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL 6 (bases 1 to 3542)

REFERENCE Fukuda, S., Furuno, M., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Adachi, J., Aizawa, K., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kageura, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, R., Saichoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

COMMENT

FEATURES

source

1. 3542
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:E130302P16"
/db_xref="taxon:10090"
/clone="E130302P16"
/tissue_type="eyeball"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
61..456
/note="unnamed protein product; hyaluronan synthase1 (MDJ|MG1:106590, GB|NM_008215, evidence: BLASTN, 99%, match=762)
putative"
/codon_start=1
/protein_id="BAC35492.1"
/db_xref="GI:26343671"
/translation="MRQDMKPKSEARCCSGLARRALTIIFALLILGLMTWAYAGVP
LASDRGLLAFLGILYAFSLVAHQSLFAYLEHRRVAAARRSLAKGLDAAATRSVA
LTISAYQEDPAYLRQCLTSARALLYPHT"

CDS

ORIGIN

Query Match 44.8%; Score 479.4; DB 3; Length 3542;
Best Local Similarity 81.6%; Pred. No. 2.7e-86;
Matches 598; Conservative 0; Mismatches 116; Indels 19; Gaps 3;

Qy 10 CAGGACGCGCCCAAGCCCACTCTCTGAGCGCCGCGCTCTCGGCCCTTGGCCCGGAGGGTG 69
|||||

Db 67 CAGGACATGCGCAAGCCCTCAGAGGAGCGGTTGCTGCTTGGCCTGGCCAGCGAGCA 126
|||||

Qy 70 CTGACCATCGCCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTACGCGCGCGG 129
|||||

Db 127 CTCACGATCATCTTGGCCTGCTCATCTGGGCTCATGACCTGGGCTTACGCGCGCAGC 186
|||||

Qy 130 GTGCGCTGGCCTCCGATCGCTACGGCTCTTGGCCTTGGCCTTACGCGCGCTTCTT 189
|||||

Db 187 GTTCCTCTGGCTTCAGATCGCTATGGACTCTTGGCCTTGGCCTTATGGGCGCATCTCTC 246
|||||

Qy 190 TCAGCGCACCCTGGTGGCGCAGAGCCTCTTTCGCTACCTGGAGCACCGCGCGCGCG 249
|||||

Db 247 AGGCGACACCTAGTGGCAGAGAGCCTCTTTCGCTTACCTGGAGCACCGAGGGTGGCAG 306
|||||

Qy 250 GCGGCGCG-----GGGCGCGCTGGATGACGCCACCGCGCGCAGTGTGGCG 294
|||||

Db 307 GCTGCGCGCGCTCTCTTGGCGAAGGGGCGCCCTGGATGCGGCCCATGACGCGCGTGGCA 366
|||||

Qy 295 CTGACCATCTCCGCTTACAGAGAGGACCGCGCTTACCTGGCCAGTCCCTGGGCTCGCC 354
|||||

Db 367 CTCACCATCTCAGCCTTACCAAGAGGATCCGCGTTTACCTGGCCAGTCTTGACCTCCGCG 426
|||||

Qy 355 CGGCGCTCTGTATACCGCGCGCGCTGCTGGGCTCTCTATGTGTGTGGTGGCAACCGC 414
|||||

Db 427 CGGCGCTTGTGTATCCCGCACAGAGTAACGCGTGTCTATGTGTGGAGCGCAACCGC 486
|||||

Qy 415 GCGGAGACCTCTTACATGTGTGACATGTTCCGCGAGGTTCTTGTGTGACGAGACCCCGCC 474
|||||

Db 487 GCTGAGGATCTGTACATGTGTGACATGTTCCGAGAGTCTTTCGCGCATGAGGACCCCGCC 546
|||||

QY 475 ACGTACGTGTGGACGCGCAACTACACAGCCCTGGGAACCGCGCGG---CGGCGCG 531
| | | | |
Db 547 ACTTATGTGTGGATGGCACTTACCATCAGCCCTGGGAACCGAGGCTACGCGCGCT 606
| | | | |
QY 532 GTGGCGCGCGAGCCTATCGGAGGTGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAG 591
| | | | |
Db 607 GTCGGTGAAGGTGCGCTACCGGAGGTGAGCGGAGGACCCCGCGCGTGGCGGTGGAG 666
| | | | |
QY 592 GCGCTGTGTGAGGAC-TCGCGAGGTGCGTGTGCGTGGCGAGCGCTGGCGCGCAAGCGGA 650
| | | | |
Db 667 GCGCTGTGTGAGACCGCAGGTGCGTGTGCGTGGCTCAGCGTTGGCGCGCAACGCTGA 726
| | | | |
QY 651 GGTGATGTACACAGCCTTCAAGGCGCTCGGAGATTCGGTGGACTACGTGCGAGTCTGTGA 710
| | | | |
Db 727 GGTGATGTACACAGCTTTCAGGCACTGGCGGACTCCGTGGACTACGTGCGAGTGA 786
| | | | |
QY 711 CTCGACACACAGG 723
| | | | |
Db 787 TGGGTTTCCACAGG 799
| | | | |

RESULT 8

BB665418
LOCUS BB665418 RIKEN full-length enriched, 0 day neonate eyeball Mus
DEFINITION musculus cDNA clone E130302P16 5', mRNA sequence.

ACCESSION BB665418

VERSION BB665418.1 GI:16499056

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS 1 (bases 1 to 650)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Hara, A., Hiramoto, K., Hori, F., Itoh, Y., Ito, M., Kawai, J.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Segabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)

TITLE

JOURNAL

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Wachiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,

Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.

and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,

Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,

Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and

Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

e mouse

FEATURES

source

Location/Qualifiers

1..650

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clones="E130302P16"

/tissue_type="eyeball"

/dev_stage="0 day neonate"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, 0 day neonate

eyeball"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN. Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'

GAGAGAGAGCGCGCGCACTCGAGTTTTTTTTTTTTTTVN 3'], cDNA was

prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by

cap-trapper. Second strand cDNA was prepared with the

primer adapter of sequence [5'

GAGAGAGAGATTCTCGAGTTAATTAATTAATCCCGCCCCCCC 3']. cDNA

was cleaved with BamHI and XhoI. Vector: a modified

pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 33.4%; Score 357.4; DB 2; Length 650;

Best Local Similarity 77.4%; Pred. No. 1e-61;

Matches 452; Conservative 0; Mismatches 117; Indels 15; Gaps 1;

QY 10 CAGGACGCGCCCAAGCCCACTCTCTGCAGCCCGCGCTCTCGGCTGCGCGGAGGTG 69

Db 67 CAGGACATGCCAAAGCCCTCAGAGCGAGCGGTGTCTCTGCGCTGCGCGAGCGAGCA 126

QY 70 CTGACCATCGCTTCGCGCTGCTCATCTCTGGGCTCATGACCTGGGCTACGCGCGCGG 129

Db 127 CTCAGATCATCTTTTGGCTGCTCATCTCTGGGCTCATGACCTGGGCTACGCGCGAGC 186

QY 130 GTGCGCTGCGCTCGCATCGCTACGGCTCTCGGCTCTCGGCTCTACGGGCTCTCTT 189

Db 187 GTTCTCTGGCTTCAGATCGCTATGAGACTCTCGGCTTTGGGCTCTATGGGCTATCTC 246

QY 190 TCAGCGCACCTGTGGCGCAGAGCCTCTTTCGCGTACCTGGAGACCGCGCGGTGGCGCG 249

Db 247 AGCGCACACCTAGTGGCAGAGAGCCTCTTTCGCTTACCTGGAGACCGGAAGGTGGCAGCG 306

QY 250 GCGCGCGG-----GGGCGCTGTGATGATGACCGACCGCGCGCAGTGTGGCG 294

Db 307 GCTGCGCGCGCTCTCTTTGGCGAAGGCGCCCTGTGATGCGGCCTCTGACGAGCGTGGCA 366

QY 295 CTGACCATCTCGGCTTACCGAGAGAGCGCGCGCTACCTGCGCCAGTCTCTGCGCTCGGCC 354

Db 367 CTCACCATCTCAGCCTTACCAAGAGATCCCGCTTACCTTGGCCAGTGTGACCTCCGCG 426

QY 355 CGCGCCCTGTGTATCCCGCGCGCGCTGCGCGCTCTCTCATGGTGGTGGATGGCAACCGC 414

Db 427 CGCGCCTTGTGTATCCCGCGCACACCGAGGTAAACGCTGCTCATGGTGGTGGACGGAACCGC 486

QY 415 GCCGAGGACCTCTACATGCTGTCGATGTTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCC 474

Db 487 GCTGAGGATCTGATGCTGAGTGTGACATGTTCCGAGAAGTCTTTCGCGCGATGAGAGACCCCGCC 546

QY 475 ACGTACGTGTGGGACGCACTACACAGCCCTGGGAACCCCGCGCGCGCGCGCGGTG 534

Db 547 ACTTATGTGGGATGGCACTACATCAGCCCTGGGACCGGGAAGTCTACAGCGCGCTG 606

QY 535 GCGCGCGAGCCTATCGGGAGGTGGAGCGGAGGATCTCTGGCGG 578


```
186 CTCTGGACTCTCGGTACGTTCTTAAAGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGG 245
875 AGCGGGCTTGTGACAGCTACTTCCACTGTGTATCTGTCATCAGCGGTTCTCTAGG 929
246 AGCGGGCTTGTGACAGCTACTTCCACTGTGTATCTGTCATCAGCGGTTCTCTAGG 300

RESULT 11
BG707608 551 bp mRNA linear EST 07-MAY-2001
LOCUS 602670672P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793704 5',
DEFINITION mRNA sequence.
ACCESSION BG707608
VERSION BG707608.1 GI:13984125
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 551)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10673 row: o column: 17
High quality sequence stop: 551.
FEATURES
Location/Qualifiers
source 1..551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4793704"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo-dt primed using primer
5'-TTTTTTTTTTTTTTN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein,
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 27.2%; Score 291.8; DB 4; Length 551;
Best Local Similarity 99.3%; Pred. No. 1.6e-48;
Matches 293; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 635 GGGCGGCAAGCGGAGTCTATGTACACAGCCTTCAAGCGCGCTCGGAGATTTCGGTGGACT 694
Db 6 GGGCGGCAAGCGGAGTCTATGTACACAGCCTTCAAGCGCGCTCGGAGATTTCGGTGGACT 65

Qy 695 ACGTGCGAGTCTGTGACTCGGACCAAGGTTGGACCCCATGCGACTGTGAGCTCGTGC 754
Db 66 ACGTGCAGGTCGTGACTCGGACCAAGGTTGGACCCCATGCGACTGTGAGCTCGTGC 125

Qy 755 GGGTACTGGACGAGACCCCGGCTAGGGCTGTGTTGGTGGGAGCTGCGGATCCTTAACC 814
Db 126 GGGTACTGGACGAGACCCCGGCTAGGGCTGTGTTGGTGGGAGCTGCGGATCCTTAACC 185
```

```
Qy 815 CTCTGGACTCTCGGTACGTTCTTAAAGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGG 874
Db 186 CTCTGGACTCTCGGTACGTTCTTAAAGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGG 245

Qy 875 AGCGGGCTTGTGACAGCTACTTCCACTGTGTATCTGTCATCAGCGGTTCTCTAGG 929
Db 246 AGCGGGCTTGTGACAGCTACTTCCACTGTGTATCTGTCATCAGCGGTTCTCTAGG 300

RESULT 12
CE377866 508 bp DNA linear GSS 27-SEP-2003
LOCUS tigr-gss-dog-17000362008906 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE377866
VERSION CE377866.1 GI:36605926
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 508)
AUTHORS Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Frazer, C.M. and
Venter, J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
Class: shotgun.
FEATURES
Location/Qualifiers
source 1..508
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 25.0%; Score 268.2; DB 9; Length 508;
Best Local Similarity 91.1%; Pred. No. 9e-44;
Matches 285; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 397 GTGTGTGATGGCAACCGCGCGAGGACCTCTACATGGTCGACATGTTCCGCGAGTCTTC 456
Db 1 GTGTGTGACGCAACCGCGCGAGGACCTCTACATGGTGGACATGTTCCGCGAGTCTTC 60

Qy 457 GCTCAGCAGGACCCCGCCACGTAGTGTGGACCGCAACTACCACCGCTTGGGAACCC 516
Db 61 GCCGACGAGGACCCCGGCGACCTAGTGTGGACCGCAACTACCACCGCTTGGAGCG 120

Qy 517 GCGCGCGCGCGCGGTGGCGCGCGAGCCCTATCGGAGGTGGAGCGCGAGGATCCTGG 576
Db 121 GCGCGCGCGCGCGCGCGCGCGCGCGCTTACCGGAGGTGGAGCGCGAAGACCCCGG 180

Qy 577 CGGTGTGAGTGGAGGCGCTGCTGAGGACTCGCAGGTGCGTGTGCTGCGCGAGCGCTGG 636
Db 181 CGGTGTGCGGTGGAGGCGCTGCTGAGGACGACAGGTGCGTGTGCTGCGCGAGCGCTGG 240

Qy 637 GCGCGCAAGCGCGAGGTTCATGTACACAGCCCTTCAAGCGCTTCGAGATTTCGTGACTAC 696
Db 241 GCGCGCAAGCGCGAGGTTCATGTACACAGCCCTTCAAGCGCTTCGAGATTTCGTGACTAC 300

Qy 697 GTGAGGTCTGTG 709
```

```
Db 301 GTGACGCTGAGTG 313
|||||
RESULT 13
BX327795 610 bp mRNA linear EST 08-APR-2004
LOCUS
DEFINITION
BX327795 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1027Y002 5-PRIME, mRNA sequence.
ACCESSION
BX327795
VERSION
BX327795.2 GI:46277978
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 610)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT
On May 2, 2003 this sequence version replaced gi:30340686.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS0BAG028ZH06_CS02640_1&c=9502.r

FEATURES
source
Location/Qualifiers
1..610
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1027Y002"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="First strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 24.2%; Score 258.8; DB 5; Length 610;
Best Local Similarity 97.0%; Pred. No. 7e-42;
Matches 263; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 659 ACACAGCGCTCGAGATTCGGTGGACTACGTGCGAGTCTGTGACTCGGACA 718
|||||
Db 1 ACATGCTTTAAGGCGCTCGAGATTCGGTGGACTACGTGCGAGTCTGTGACTCGGACA 60
|||||
Qy 719 CAAGTTGGACCCCATGGCACTGTGTGAGTCTGTGCGGGTACTGGAACGAGACCCCGGG 778
|||||
Db 61 CAAGTTGGACCCCATGGCACTGTGTGAGTCTGTGCGGGTACTGGAACGAGACCCCGGG 120
|||||
Qy 779 TAGGGGCTGTGGTGGGAGCTGGGATCTTAACTCCCTCTGGACTCTGGGTGAGCTTCC 838
|||||
Db 121 TAGGGGCTGTGGTGGGAGCTGGGATCTTAACTCCCTCTGGACTCTGGGTGAGCTTCC 180
|||||
Qy 839 TAAGCAGCCTCGGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCC 898
|||||
Db 181 TAAGCAGCCTCGGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCC 240
|||||
Qy 899 ACTGTGTATCTGATCAGCGGTTCTCTAGG 929
|||||
Db 241 ACTGTGTATCTGATCAGCGGTTCTCTAGG 271
|||||

RESULT 14
CF239758 894 bp mRNA linear EST 05-AUG-2003
AGENCOURT 15100464 NICHG_XGC_Emb6 Xenopus tropicalis cDNA clone
IMAGE:6994934 5', mRNA sequence.
CF239758
ACCESSION
CF239758.1 GI:33442966
VERSION
CF239758.1
KEYWORDS
EST.
SOURCE
Xenopus tropicalis (western clawed frog)
ORGANISM
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 894)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM14677 row: m column: 13
High quality sequence stop: 699.

FEATURES
source
Location/Qualifiers
1..894
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6994934"
/tissue_type="neurala"
/dev_stage="embryo, stages 14-19"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHG_XGC Emb6"
/notes="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.1 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

ORIGIN
Query Match 23.9%; Score 256.2; DB 7; Length 894;
Best Local Similarity 63.6%; Pred. No. 2.3e-41;
Matches 409; Conservative 0; Mismatches 228; Indels 6; Gaps 1;
Qy 289 GTGGCGGTGACCATCTCCGCTTACAGGAGAGACCCCGCGTACCTGCGCGAGTGGCTGCG 348
Db 54 GTGGCGGTGACCATCTCCGCTTACAGGAGAGACCCCGCGTACCTGCGCGAGTGGCTGCG 113
|||||
Qy 349 TCCGCGCGCGCTGCTGTACCCGCGCGCGGCTGCGCTCATGATGCTGCTGCTGCTGCTG 408
|||||
Db 114 TCCTGCGGTGACGTGAAGTACCCCAAGGACAAACTCAAGATCATTTTGGTCATCGACGG 173
|||||
Qy 409 AACCGCGCCGAGGACCTCTACATGCTCGACATGTTCCGCGAGGCTTTCGCTGACGAGGAC 468
|||||
Db 174 AACAGGAGATGACCGCTACATGATGATGATGATGATGATGATGATGATGATGATGATG 233
|||||
Qy 469 CCCGCCACGTGCTGTGGGACGGCACTACACAGCCCTGGGACCCCGCGCGCGCGCGG 528
Db 234 GTGGGTACCTACCTCTGTGGAAGGGCACTACACACTGGCGTAAAGAGAGACCAAGATG 293
|||||
Qy 529 GCGGTGGCGCGGAGCTATCGGAGGTGAGCGGAGGATCTCTGGCGCGCTGGCAGTG 588
|||||
Db 294 TCCTGTCTGAGGTGTCGAAGCCCTGAATGAAGATGAG-----GGGATACGTATAGTG 347
|||||
Qy 589 GAGGCGCTGTGAGGACTCGCAGGTGCTGTGCTGGCGAGCGCTCGCGGCGGCAAGCGC 648
|||||
```

Db 348 GAGGAACATGGTCAGAACCAAGAGGTGTGTGTGCATCATGCAGCACTGGGGCGGCAACGG 407
Qy 649 GAGGTCAATGTATACAGCCCTTCAAGCGCTTCGAGAGATTCGGTGGACTAGCTGCAGGTCTGT 708
Db 408 GAGGTCAATGTATACAGCCCTTTCGGGCCATTCGGGACCAACATGAGTACGTACAGTCTGT 467
Qy 709 GACTCGGACACAGGTTTGGACCCCATGCGACTGTCTGGAGCTCGTCCGGGTACTGACACGAG 768
Db 468 GACTCCGACACCAACTCGGACCACTGGCCACGGTGGAAATGGTGAAGGTCTCTAGAGGCC 527
Qy 769 GACCCCGGGTAGGGGCTGTGTGTGGGACGTGGGATCCCTTAACCCCTCTGGACTCCCTGG 828
Db 528 AACGAGCTGTGGGGGCGGTGGAGAGACGTTCGCATCCCTGAACCCCTAGACCTCTTC 587
Qy 829 GTCAGCTTCCTAAGCAGCTCGATACATGGGTAGCTTCAATGTGGAGCGGGCTTGTGAG 888
Db 588 ATCAGTTTCATGAGCAGCTAGCTTATTTGGATGGGTTTAACTGGAGAGGCGCTGCCAG 647
Qy 889 AGCTACTTCCACTGTGTATCTTCGATCAGCGGTTCCTTAGGTA 931
Db 648 TCCTACTTCGACTGCGTCTCTGTATCAGTGGCCCCCTGGGAA 690

RESULT 15

BI753116 782 bp mRNA linear EST 25-SEP-2001
LOCUS 603025961F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196399 5',
DEFINITION mRNA sequence.

ACCESSION BI753116

VERSION BI753116.1 GI:15744694

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 782)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L14M11491 row: j column: 16

High quality sequence stop: 780.

Location/Qualifiers

FEATURES

Source

1. 782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5196399"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 23.9%; Score 255.8; DB 4; Length 782;
Best Local Similarity 99.2%; Pred. No. 2.8e-41;
Matches 257; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 671 AGGCGCTCGGAGATTCCGTGGACTACGTGCAGGTCTGTGACTCGGACACAAGGTTGGACC 730
Db 6 AAGCGCTCGGAGATTCCGTGGACTACGTGCAGGTCTGTGACTCGGACACAAGGTTGGACC 65
Qy 731 CCATGGCACTGCTGGAGCTCGTGGGGTACTGGACGAGACCCCGGGTAGGGGCTGTG 790
Db 66 CCATGGCACTGCTGGAGCTCGTGGGGTACTGGACGAGACCCCGGGTAGGGGCTGTG 125
Qy 791 GTGGGGACGTGCGGATCCCTTAACCCCTCTGGACTCCTGGGTGAGCTTCTTAAGCAGCCTGC 850
Db 126 GTGGGGACGTGCGGATCCCTTAACCCCTCTGGACTCCTGGGTGAGCTTCTTAAGCAGCCTGC 185
Qy 851 GATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCACTGTGTATCCT 910
Db 186 GATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCACTGTGTATCCT 245
Qy 911 GCATCAGCGGTTCCTTAGG 929
Db 246 GCATCAGCGGTTCCTTAGG 264

Search completed: March 13, 2005, 11:54:55
Job time : 4422.94 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 01:47:21 ; Search time 696.034 Seconds
(without alignments)
9108.809 Million cell updates/sec

Title: US-10-672-399-3
Perfect score: 1071
Sequence: 1 atgagacagcaggacgccc.....ccggaccatgcgtggatga 1071

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980e:*
2: Geneseqn1990e:*
3: Geneseqn2000e:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	927.4	86.6	1737	10 ADL13691	ADL13691 Osteoarth
2	927.4	86.6	2088	10 ADL13690	ADL13690 Osteoarth
3	927.4	86.6	2116	2 AAT99541	AAT99541 Human hya
4	925.8	86.4	1737	10 ADL13692	ADL13692 Osteoarth
5	925.8	86.4	2087	10 ADL13694	ADL13694 Osteoarth
6	905.4	84.5	2117	2 AAT96713	AAT96713 Human hya
7	697.2	65.1	231222	10 ADL13693	ADL13693 Osteoarth
8	678.2	63.3	1752	8 ABZ76734	ABZ76734 Mouse hya
9	678.2	63.3	1752	10 AAD59442	AAD59442 Mouse hya
10	678.2	63.3	2102	2 AAT91655	AAT91655 Mouse hya
11	678.2	63.3	2102	2 AAZ10862	AAZ10862 Hyaluron
12	678.2	63.3	2102	3 AAZ88199	AAZ88199 Mouse hya
13	678.2	63.3	2102	3 AAZ39987	AAZ39987 Murine HA
14	422	39.4	662	6 ABQ27384	ABQ27384 Oligonuc1
15	422	39.4	662	6 ABQ27385	ABQ27385 Oligonuc1
16	393.8	36.8	662	6 ABQ27382	ABQ27382 Oligonuc1
17	393.8	36.8	662	6 ABQ27383	ABQ27383 Oligonuc1
18	338.6	31.6	490	9 ACH39998	ACH39998 Human toe
19	290.6	27.1	1783	10 ADC49212	ADC49212 Rabbit hy
20	272.8	25.5	1662	8 ACC51023	ACC51023 Human bla

21	272.8	25.5	1662	8 ABX76415	ABX76415 Lung canc
22	272.8	25.5	1662	11 ADN39051	ADN39051 Cancer/an
23	272.8	25.5	4049	8 ABX76214	ABX76214 Lung canc
24	272.8	25.5	4049	12 ADN05115	ADN05115 Antipsori
25	272.8	25.5	4220	12 ADJ74905	ADJ74905 Marker ge
26	248.4	23.2	1767	10 ADD93927	ADD93927 Xenopus 1
27	236.2	22.1	1665	2 AAV18822	AAV18822 Murine hy
28	236.2	22.1	1665	3 AAZ88201	AAZ88201 Mouse hya
29	236.2	22.1	1665	8 ABZ76736	ABZ76736 Mouse hya
30	236.2	22.1	1665	10 AAD59444	AAD59444 Mouse hya
31	236.2	22.1	5919	12 ADJ75814	ADJ75814 Marker ge
32	207.4	19.4	1229	12 ADJ74906	ADJ74906 Marker ge
33	176.2	16.5	1653	6 ABL41013	ABL41013 Murine hy
34	176.2	16.5	1659	8 ABZ76735	ABZ76735 Mouse hya
35	176.2	16.5	1659	10 AAD59443	AAD59443 Mouse hya
36	176.2	16.5	2947	2 AAV18821	AAV18821 Murine hy
37	176.2	16.5	2948	6 ABI98697	ABI98697 Mouse isc
38	176.2	16.5	4194	9 ADA49685	ADA49685 DNA encod
39	176.2	16.5	4194	10 ADC59298	ADC59298 Mouse HAS
40	162.6	15.2	2890	3 AAZ88200	AAZ88200 Mouse hya
41	148.8	13.9	1659	10 ADL13696	ADL13696 Osteoarth
42	148.8	13.9	3003	6 ABK84373	ABK84373 Human CDN
43	148.8	13.9	3003	9 ADA49683	ADA49683 DNA encod
44	148.8	13.9	3003	10 ADC59296	ADC59296 Human HAS
45	148.8	13.9	3003	10 ADD12718	ADD12718 Human CDN

ALIGNMENTS

RESULT 1
ADL13691
ID ADL13691 standard; DNA; 1737 BP.
XX
AC ADL13691;
XX
DT 06-MAY-2004 (first entry)
XX
DE Osteoarthritis-associated polymorphic nucleotide #223.
XX
KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO2003054166-A2.
XX
PD 03-JUL-2003.
XX
PF 19-DEC-2002; 2002WO-US041225.
XX
PR 20-DEC-2001; 2001US-0342603P.
XX
(INCY-) INCYTE GENOMICS INC.
XX
Jones KA, Schafer A;
XX
DR WPI; 2003-559141/52.
XX
PT Determining susceptibility of an individual to joint space narrowing,
PT osteophyte development and/or joint pain comprises identifying whether
PT the individual has at least one polymorphism in a polynucleotide encoding
PT a protein.
XX
PS Disclosure; SEQ ID NO 223; 297pp; English.
XX
CC The invention relates to a method of determining susceptibility of an
CC individual to joint space narrowing and/or osteophyte development and/or
CC joint pain comprising identifying whether the individual has at least one
CC polymorphism in a polynucleotide encoding at least one of the protein
CC listed in the specification. The methods, composition and agent are
CC useful for modulating the susceptibility of an individual to joint space

CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).

XX SQ Sequence 1737 BP; 232 A; 571 C; 603 G; 331 T; 0 U; 0 Other;

Query Match 86.6%; Score 927.4; DB 10; Length 1737;
Best Local Similarity 99.9%; Pred. No. 1.6e-152;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCACTCTCTGACGACCGCCGCTGCTCGGCTTGCC 60
Db 1 ATGAGACAGCAGGACGCGCCCAAGCCACTCTCTGACGACCGCCGCTGCTCGGCTTGCC 60
Qy 61 CGAGGGTGTGTAACCATCGCTTGGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 120
Db 61 CGAGGGTGTGTAACCATCGCTTGGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 120
Qy 121 GCCGCCGGGGTGGCGCTGGCTCCGATCGCTACGGCTCTCTGGCTTGGGCTCTACGGG 180
Db 121 GCCGCCGGGGTGGCGCTGGCTCCGATCGCTACGGCTCTCTGGCTTGGGCTCTACGGG 180
Qy 181 GCCTTCCTTTTCAGCGCACCTGTGTGCGCAGAGCCTCTTCTGGCTGCTCTGAGCAGCGGCGG 240
Db 181 GCCTTCCTTTTCAGCGCACCTGTGTGCGCAGAGCCTCTTCTGGCTGCTCTGAGCAGCGGCGG 240
Qy 241 GTGGCGGCGGCGCGCGCGCGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 241 GTGGCGGCGGCGCGCGCGCGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 300
Qy 301 ATCTCCGCTTACAGGAGGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 ATCTCCGCTTACAGGAGGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Qy 361 CTGCTGTATCCCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 CTGCTGTATCCCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy 421 GACCTTATATGCTGACATGTTTCGCGAGGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 GACCTTATATGCTGACATGTTTCGCGAGGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 481 GTGTGGGACGCAACTACCAACGCGCTGCGAACCCTGCGGAGCCCGCGCGCGCGCGCGCGCGCG 540
Db 481 GTGTGGGACGCAACTACCAACGCGCTGCGAACCCTGCGGAGCCCGCGCGCGCGCGCGCGCGCG 540
Qy 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGGTG 600
Db 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGGTG 600
Qy 601 AGGACTCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 AGGACTCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 661 ACAGCCTTCAAGGCGCTCGAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 ACAGCCTTCAAGGCGCTCGAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 721 AGTTTGAACCCATGCGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 AGTTTGAACCCATGCGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 GGGGCTGTGTGGGACGCTGCGATCCTTAACCTCTGGAATCTCTGGGTGAGTCTGCTGCTGCTGCTGCT 840
Db 781 GGGGCTGTGTGGGACGCTGCGATCCTTAACCTCTGGAATCTCTGGGTGAGTCTGCTGCTGCTGCTGCT 840
Qy 841 AGCAGCTGCGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCAC 900

Db 841 AGCAGCTGCGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCAC 900
Qy 901 TGTGTATCTCTGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 929
Db 901 TGTGTATCTCTGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 929

RESULT 2
ADL13690
ID ADL13690 standard; DNA; 2088 BP.
XX AC ADL13690;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #222.
XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN W02003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX WPI; 2003-559141/52.

Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.

Disclosure; SEQ ID NO 222; 297pp; English.

The invention relates to a method of determining susceptibility of an
individual to joint space narrowing and/or osteophyte development and/or
joint pain comprising identifying whether the individual has at least one
polymorphism in a polynucleotide encoding at least one of the protein
listed in the specification. The methods, composition and agent are
useful for modulating the susceptibility of an individual to joint space
narrowing and/or osteophyte development and/or joint pain that is
associated with a disease, preferably osteoarthritis. The cell line and
the non-human animal are useful for screening for an agent for diagnosing
an individual having susceptibility to joint space narrowing and/or
osteophyte development and/or joint pain. This sequence corresponds to
the polynucleotide encoding a protein listed in the specification. (Note:
The sequence data for this patent did not form part of the printed
specification but was obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences).

Query Match 86.6%; Score 927.4; DB 10; Length 2088;
Best Local Similarity 99.9%; Pred. No. 1.6e-152;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCACTCTCTGACGACCGCCGCTGCTCGGCTTGCC 60
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCACTCTCTGACGACCGCCGCTGCTCGGCTTGCC 95
Qy 61 CGAGGGTGTGTAACCATCGCTTGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120

```
Db 96 CGAGGGTGTGACCATGCCCTTCGCCCTGCTCATCTCGGGCTCATGACCTGGGCGCTTAC 155
Qy 121 GCCCGGGGTGCGCTGGCTCCGATCGCTACGGCTCTTGGCTCTTACGGG 180
Db 156 GCCCGGGGTGCGCTGGCTCCGATCGCTACGGCTCTTGGCTCTTACGGG 215
Qy 181 GCCTTCCTTTAGCGACCTGGTGGCGAGAGCTTTCGGGTACTTGAGGACCGGGGG 240
Db 216 GCCTTCCTTTAGCGACCTGGTGGCGAGAGCTTTCGGGTACTTGAGGACCGGGGG 275
Qy 241 GTGGCGGGCGGGCGGGCGGGCTGGATGACGACCGCGCGAGTGTGGCGTAC 300
Db 276 GTGGCGGGCGGGCGGGCGGGCTGGATGACGACCGCGCGAGTGTGGCGTAC 335
Qy 301 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGTGCTGGCGCGCGGCC 360
Db 336 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGTGCTGGCGCGGCC 395
Qy 361 CTGCTGTACCGCGCGCGGGCTGGCTGCTCATGCTGCTGATGGATGGCAACCGCGCCGAG 420
Db 396 CTGCTGTACCGCGCGCGGGCTGGCTGCTCATGCTGCTGATGGATGGCAACCGCGCCGAG 455
Qy 421 GACCTCTACATGCTGACATGCTTCGCGAGGTCTTCGCTGACGAGACCCCGCCAGTAC 480
Db 456 GACCTCTACATGCTGACATGCTTCGCGAGGTCTTCGCTGACGAGACCCCGCCAGTAC 515
Qy 481 GTGGGAGCGCACTACACAGCCCTGGGAACCCCGGGCGGGCGGGCTGGGCGCC 540
Db 516 GTGGGAGCGCACTACACAGCCCTGGGAACCCCGGGCGGGCGGGCTGGGCGCC 575
Qy 541 GGAGCCTATCGGAGGTGGAGCGAGATCTTGGCGGGCTGGCAGTGGAGCGCTGTGTG 600
Db 576 GGAGCCTATCGGAGGTGGAGCGAGATCTTGGCGGGCTGGCAGTGGAGCGCTGTGTG 635
Qy 601 AGGACTCCAGGTGGTGTGGTGGCGAGCGCTGGGCGGGCGGGCGGGCTGGGCGCTATGAT 660
Db 636 AGGACTCCAGGTGGTGTGGTGGCGAGCGCTGGGCGGGCGGGCGGGCTGGGCGCTATGAT 695
Qy 661 ACAGCTTTCAAGGCGCTCGGAGATTCGGTGGACTACGTGAGTCTGTGACTCGGACACA 720
Db 696 ACAGCTTTCAAGGCGCTCGGAGATTCGGTGGACTACGTGAGTCTGTGACTCGGACACA 755
Qy 721 AGGTTGAGACCCATGGCATGCTGGAGTCTGTGGGTACTGGAGCGAGACCCCGGGTA 780
Db 756 AGGTTGAGACCCATGGCATGCTGGAGTCTGTGGGTACTGGAGCGAGACCCCGGGTA 815
Qy 781 GGGGCTGTGGTGGGAGCTGGGATCTTAACTCTGGACTCTGGCTGGGCTGGCTTCTTA 840
Db 816 GGGGCTGTGGTGGGAGCTGGGATCTTAACTCTGGACTCTGGCTGGGCTGGCTTCTTA 875
Qy 841 AGCAGCTGCGATACTGGGTAGCCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCAC 900
Db 876 AGCAGCTGCGATACTGGGTAGCCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCAC 935
Qy 901 TGTGTATCTGCATCAGCGGTTCCTAGG 929
Db 936 TGTGTATCTGCATCAGCGGTTCCTAGG 964
```

RESULT 3

AAT99541

ID AAT99541 standard; cDNA; 2116 BP.

AC AAT99541;

XX

XX 21-MAY-1998 (first entry)

XX

XX Human hyaluronan synthase cDNA clone 30C.

XX

KW Hyaluronan synthase; HAS gene; human; hyaluronan acid; cell adhesion;

KW wound healing; vulnary; tissue repair; scar; keloid; therapy; ds.

XX

OS Homo sapiens.

```
XX Key Location/Qualifiers
FH CDS 36..1772
FT /*tag= a
FT polyA_signal 2066..2071
FT /*tag= b
XX
PN WO9740174-A1.
XX
XX 30-OCT-1997.
XX
XX 17-APR-1997; 97WO-US006350.
XX
XX 22-APR-1996; 96US-00635552.
XX
XX (LEUK-) LEUKOSITE INC.
XX
XX Briskin MJ;
XX
XX WPI; 1997-549359/50.
XX
XX P-PSDB; AAW26765.
XX
XX Human hyaluronan synthase - useful for recombinant production of
XX hyaluronic acid for wound healing, tissue repair and reducing
XX hypertrophic scar and keloid formation.
XX
XX Claim 3; Page 36-38; 58pp; English.
XX
XX cDNA clone 30C includes a coding region for hyaluronan synthase (HAS)
XX (see AAW26765), an enzyme involved in the synthesis of hyaluronan
XX (hyaluronic acid) and which has the ability to confer cell adhesion by
XX the lymphocyte receptor CD44. Clone 30C was isolated using an expression
XX cloning system developed to isolate cDNA clones that encode proteins that
XX confer adhesion of the murine T cell lymphoma TK1. A human mesenteric
XX lymph node expression library was constructed that, upon transfection
XX into CHO/P cells, yielded clone 30C that mediated rosetting of TK1 cells
XX to some of the transfectants. The isolated clone can be utilised in a
XX claimed method for producing HAS in host cells. Such host cells are used
XX in a claimed method for the production of hyaluronan. Hyaluronan is
XX useful for wound healing and tissue repair, and can reduce or prevent
XX hypertrophic scars and keloid formation. It is also used in eye surgery
XX as a replacement for vitreous fluid
XX
XX Sequence 2116 BP; 330 A; 651 C; 712 G; 423 T; 0 U; 0 Other;
```

```
Query Match 85.6%; Score 927.4; DB 2; Length 2116;
Best Local Similarity 99.9%; Pred. No. 1.6e-152;
Matches 528; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAGACAGCAGGACGGCGCCCAAGCCCACTCTGACAGCCCGCGCTGCTCGGCTGGCC 60
Db 36 ATGAGACAGCAGGACGGCGCCCAAGCCCACTCTGACAGCCCGCGCTGCTCGGCTGGCC 95
Qy 61 CGGAGGTGTGTGACATCGCCTTCGCCCTGTCTATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 96 CGGAGGTGTGTGACATCGCCTTCGCCCTGTCTATCTCTGGGCTCATGACCTGGGCTTAC 155
Qy 121 GCCCGCGGGTGGCGCTGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTTAC 180
Db 156 GCCCGCGGGTGGCGCTGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTTAC 215
Qy 181 GCCTTCCTTTTTCAGCGCACCTGCTGGCGCGCAGAGCTCTTTCGGCTTACCTGGAGCACCGGG 240
Db 216 GCCTTCCTTTTTCAGCGCACCTGCTGGCGCGCAGAGCTCTTTCGGCTTACCTGGAGCACCGGG 275
Qy 241 GTGGCGGGCGGGCGGGCGGGCGGGCTGGATGACGACCGCGCGAGTGTGGCGTAC 300
Db 276 GTGGCGGGCGGGCGGGCGGGCGGGCTGGATGACGACCGCGCGAGTGTGGCGTAC 335
Qy 301 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGTGCTGGCGCGGCC 360
Db 336 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGTGCTGGCGCGGCC 395
```

361 CTGCTGTACCCGCGCGCGCTCGCGTCTCATGTGTGGATGGCAACCGCGCGAG 420
|||||
396 CTGCTGTACCCGCGCGCGCGCTCGCGTCTCATGTGTGGATGGCAACCGCGCGAG 455
|||||
421 GACCTCTACATGGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC 480
456 GACCTCTACATGGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC 515
|||||
481 GTGTGGGACGCACTACACAGCCCTGGGAACCCCGCGCGCGCGCGCGGTGGCGCC 540
516 GTGTGGGACGCACTACACAGCCCTGGGAACCCCGCGCGCGCGCGCGGTGGCGCC 575
|||||
541 GGAGCTATCGGAGGTGGAGCGGAGATCTCGGCGGCTGGCAGTGGAGGCGCTGTGTG 600
576 GGAGCTATCGGAGGTGGAGCGGAGATCTCGGCGGCTGGCAGTGGAGGCGCTGTGTG 635
|||||
601 AGGACTCGCAGGTGCTGTGGTGGCGAGCGCTGGGCGCGCAAGCGCGAGGTTCATGTAC 660
636 AGGACTCGCAGGTGCTGTGGTGGCGAGCGCTGGGCGCGCAAGCGCGAGGTTCATGTAC 695
|||||
661 ACAGCTTCAAGCGCTCGGAGATTCGGTGAAGTCTGTGACTCGAGTCTGTGACTCGGACACA 720
696 ACAGCTTCAAGCGCTCGGAGATTCGGTGAAGTCTGTGACTCGAGTCTGTGACTCGGACACA 755
|||||
721 AGGTGGACCCCATGGCACTCTGAGCTCTGCGGGTACTGGAGAGAGCCCGCGGTA 780
756 AGGTGGACCCCATGGCACTCTGAGCTCTGCGGGTACTGGAGAGAGCCCGCGGTA 815
|||||
781 GGGGCTGTGTGGGAGCGTGGGATCCTTAACCTCTGGAAGTCTGTGAGTCTGTGACTCGGTA 840
816 GGGGCTGTGTGGGAGCGTGGGATCCTTAACCTCTGGAAGTCTGTGAGTCTGTGACTCGGTA 875
|||||
841 AGCAGCTGCGATGACTGGGTAGCTTCAATGTGGAGCGGGTGTGAGAGTACTTTCAC 900
876 AGCAGCTGCGATGACTGGGTAGCTTCAATGTGGAGCGGGTGTGAGAGTACTTTCAC 935
|||||
901 TGTGTATCTCATCAGCGGTCTCTAGG 929
936 TGTGTATCTCATCAGCGGTCTCTAGG 964

RESULT 4

ID ADL13692 standard; DNA; 1737 BP.

XX AC ADL13692;

XX DT 06-MAY-2004 (first entry)

XX DE Osteoarthritis-associated polymorphic nucleotide #224.

XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX joint space narrowing; osteophyte development; joint pain;
XX osteoarthritis; SNP; single nucleotide polymorphism.

XX OS Homo sapiens.

XX PN WO2003054166-A2.

XX PD 03-JUL-2003.

XX PF 19-DEC-2002; 2002WO-US041225.

XX PR 20-DEC-2001; 2001US-0342603P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Jones KA, Schafer A;

XX DR WPI; 2003-559141/52.

XX PT Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether

PT the individual has at least one polymorphism in a polynucleotide encoding
PT a protein.

XX PS Disclosure; SEQ ID NO 224; 297pp; English.

XX CC The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polynucleotide encoding a protein listed in the specification. (Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences).

SQ Sequence 1737 BP; 232 A; 569 C; 604 G; 332 T; 0 U; 0 Other;

Query Match 86.4%; Score 925.8; DB 10; Length 1737;
Best Local Similarity 99.8%; Pred. No. 3.1e-152;
Matches 927; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAGACAGCAGGACGCGCCAAAGCCCACTCTGACGCGCGCGCTGCTCGGCGCTGGCC 60
Db 1 ATGAGACAGCAGGACGCGCCAAAGCCCACTCTGACGCGCGCGCTGCTCGGCGCTGGCC 60
Qy 61 CGGAGGGTGTGACCATCGGCTTCGGCTCTCATCTCGGCGCTCATGACCTGGCGCTTAC 120
Db 61 CGGAGGGTGTGACCATCGGCTTCGGCTCTCATCTCGGCGCTCATGACCTGGCGCTTAC 120
Qy 121 GCGCGCGGGTGCCTGCGCTTCGATCGCTACGCGCTCTCGGCTTCGCGCTCTACGGG 180
Db 121 GCGCGCGGGTGCCTGCGCTTCGATCGCTACGCGCTCTCGGCTTCGCGCTCTACGGG 180
Qy 181 GCCTTCCTTTCAGCGCACTGTGGCGCAGAGCTCTTCGGGTACCTGGAGCAGCGCGG 240
Db 181 GCCTTCCTTTCAGCGCACTGTGGCGCAGAGCTCTTCGGGTACCTGGAGCAGCGCGG 240
Qy 241 GTGCGCGCGCGCGCGCGCGCGCTGATGACGACCGCGCGCTGTGGCGCTGACC 300
Db 241 GTGCGCGCGCGCGCGCGCGCGCTGATGACGACCGCGCGCTGTGGCGCTGACC 300
Qy 301 ATCTTCGCGCTTACCAGGAGGACCCCGCGTACCTGCGCCAGTTCCTGGCGTTCGCGCGCC 360
Db 301 ATCTTCGCGCTTACCAGGAGGACCCCGCGTACCTGCGCCAGTTCCTGGCGTTCGCGCGCC 360
Qy 361 CTGCTGTACCGCGCGCGCGCTGCGGCTCTCATGTGTGGATGGCAACCGCGCGGAG 420
Db 361 CTGCTGTACCGCGCGCGCGCTGCGGCTCTCATGTGTGGATGGCAACCGCGCGGAG 420
Qy 421 GACCTCTACATGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC 480
Qy 481 GTGTGGGACGCAACTACCAACCGCTTGGGAACCCCGCGCGCGCGCGGTGGCGGCC 540
Db 481 GTGTGGGACGCAACTACCAACCGCTTGGGAACCCCGCGCGCGCGCGGTGGCGGCC 540
Qy 541 GGAGCTTATCGGAGGTGGAGCGGAGATCTCGGCGGCTGGCAGTGGAGGCGCTGTGTG 600
Db 541 GGAGCTTATCGGAGGTGGAGCGGAGATCTCGGCGGCTGGCAGTGGAGGCGCTGTGTG 600
Qy 601 AGGACTCGCAGGTGCTGTGGCGCAGCGCTGGGCGCGCAAGCGCGAGGTTCATGTAC 660
Db 601 AGGACTCGCAGGTGCTGTGGCGCAGCGCTGGGCGCGCAAGCGCGAGGTTCATGTAC 660
Qy 661 ACAGCTTCAAGCGCTTCGGGATTCGGTGGAGTACCTGAGAGTCTGTGACTCGGACACA 720
Db 661 ACAGCTTCAAGCGCTTCGGGATTCGGTGGAGTACCTGAGAGTCTGTGACTCGGACACA 720

Db 661 ACAGCCTTCAGGCGCTCGGAGATTCTGGTGGACTACGTGCGAGGTCTGTGACTCGGACACA 720
QY 721 AGGTTGGACCCCATGGCACTGCTGAGCTCGTGGCGGTACTGGACGAGACCCCGGGTA 780
Db 721 AGGTTGGACCCCATGGCACTGCTGAGCTCGTGGCGGTACTGGACGAGACCCCGGGTA 780
QY 781 GGGGCTGTTGGTGGGAGCGTGGGATCCTTAACTCTGGACTCTGGGTGAGTTCTTA 840
Db 781 GGGGCTGTTGGTGGGAGCGTGGGATCCTTAACTCTGGACTCTGGGTGAGTTCTTA 840
QY 841 AGCAGCCTGGGATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
Db 841 AGCAGCCTGGGATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
QY 901 TGTGTATCTGCATCAGCGGTTCTCTAGG 929
Db 901 TGTGTATCTGCATCAGCGGTTCTCTAGG 929

RESULT 5
ADL13694

ID ADL13694 standard; DNA; 2087 BP.

XX XX

AC ADL13694;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #226.

XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;

KW joint space narrowing; osteophyte development; joint pain;

KW osteoarthritis; SNP; single nucleotide polymorphism.

XX Homo sapiens.

XX WO2003054166-A2.

XX 03-JUL-2003.

XX 19-DEC-2002; 2002WO-US041225.

XX 20-DEC-2001; 2001US-0342603P.

XX (INCY-) INCYTE GENOMICS INC.

XX Jones KA, Schafer A;

XX WPI; 2003-559141/52.

XX Determining susceptibility of an individual to joint space narrowing, osteophyte development and/or joint pain comprises identifying whether the individual has at least one polymorphism in a polynucleotide encoding a protein.

XX Disclosure; SEQ ID NO 226; 297pp; English.

XX The invention relates to a method of determining susceptibility of an individual to joint space narrowing and/or osteophyte development and/or joint pain comprising identifying whether the individual has at least one polymorphism in a polynucleotide encoding at least one of the protein listed in the specification. The methods, composition and agent are useful for modulating the susceptibility of an individual to joint space narrowing and/or osteophyte development and/or joint pain that is associated with a disease, preferably osteoarthritis. The cell line and the non-human animal are useful for screening for an agent for diagnosing an individual having susceptibility to joint space narrowing and/or osteophyte development and/or joint pain. This sequence corresponds to the polynucleotide encoding a protein listed in the specification. (Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences).

XX Sequence 2087 BP; 302 A; 650 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 86.4%; Score 925.8; DB 10; Length 2087;
Best Local Similarity 99.8%; Pred. No. 3.1e-152;
Matches 927; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTGTGAGCCCGCGTGTCTGGGCTTGCC 60
Db 36 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTGTGAGCCCGCGTGTCTGGGCTTGCC 95
QY 61 CGGAGGGTGTGACCATCGCCTTCGCCCTGTCACTCTGTGGCCCTCATGACCTGGGCTTAC 120
Db 96 CGGAGGGTGTGACCATCGCCTTCGCCCTGTCACTCTGTGGCCCTCATGACCTGGGCTTAC 155
QY 121 GCCCGCGGGGTGCCGCTCCGATCGCTACCGCCCTCTCGGCTTCTGGGCTTCTACGG 180
Db 156 GCCCGCGGGGTGCCGCTCCGATCGCTACCGCCCTCTCGGCTTCTGGGCTTCTACGG 215
QY 181 GCCTTCTTTCAGCGCACCTGTGGGCGCAGAGCTCTTCGGGTACTTGGAGACCGGGCG 240
Db 216 GCCTTCTTTCAGCGCACCTGTGGGCGCAGAGCTCTTCGGGTACTTGGAGACCGGGCG 275
QY 241 GTGGCGCGCGCGCGCGCGCGCTGTGATCGACGACCGCGCGCAGTGTGGCGCTGACC 300
Db 276 GTGGCGCGCGCGCGCGCGCGCTGTGATCGACGACCGCGCGCAGTGTGGCGCTGACC 335
QY 301 ATCTCCGCTTACCAGGAGGACCCCGCGTACTCTGCGCCAGTCCCTGGGCTCCCGCGGCC 360
Db 336 ATCTCCGCTTACCAGGAGGACCCCGCGTACTCTGCGCCAGTCCCTGGGCTCCCGCGGCC 395
QY 361 CTGCTGTATCCCGCGCGCGCGCTCGGCTCTCTATGTGTGGTATGGCAACCGCGCCGAG 420
Db 396 CTGCTGTATCCCGCGCGCGCGCTCGGCTCTCTATGTGTGGTATGGCAACCGCGCCGAG 455
QY 421 GACCTCTACATGTGCGACATGTTCGCGAGAGTCTTCGCTGACGAGGACCCCGCACGTCAC 480
Db 456 GACCTCTACATGTGCGACATGTTCGCGAGAGTCTTCGCTGACGAGGACCCCGCACGTCAC 515
QY 481 GTGTGGGACGGCAACTACCAAGCCCTTGGGAAACCCCGCGCGCGCGCGCGTGGGCGCC 540
Db 516 GTGTGGGACGGCAACTACCAAGCCCTTGGGAAACCCCGCGCGCGCGCGTGGGCGCC 575
QY 541 GGAGCCTATCGGGAGGTGGAGCGGAGGATCCTTGGGCGGCTGGCAGTGGAGGCGCTGTG 600
Db 576 GGAGCCTATCGGGAGGTGGAGCGGAGGATCCTTGGGCGGCTGGCAGTGGAGGCGCTGTG 635
QY 601 AGGACTCGCAGGTGCGTGTGCGCGCAGCGCTGGGGCGGCAAGCGGAGGTCATGTAC 660
Db 636 AGGACTCGCAGGTGCGTGTGCGCGCAGCGCTGGGGCGGCAAGCGGAGGTCATGTAC 695
QY 661 ACAGCCTTCAAGGCGCTCGGAGATTTCGGTGGACTACGTGACAGGTCTGTGACTCGGACACA 720
Db 696 ACAGCCTTCAAGGCGCTCGGAGATTTCGGTGGACTACGTGACAGGTCTGTGACTCGGACACA 755
QY 721 AGGTTGAGACCCCATGGCACTGCTGAGCTCTGGGGTACTGTGAGAGGACCCCGGGTA 780
Db 756 AGGTTGAGACCCCATGGCACTGCTGAGCTCTGGGGTACTGTGAGAGGACCCCGGGTA 815
QY 781 GGGGCTGTTGGTGGGAGCGTGGGATCCTTAACCTCTGGACTCTGGGTGAGTCTTCTTA 840
Db 816 GGGGCTGTTGGTGGGAGCGTGGGATCCTTAACCTCTGGACTCTGGGTGAGTCTTCTTA 875
QY 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
Db 876 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 935
QY 901 TGTGTATCTGCATCAGCGGTTCTCTAGG 929
Db 936 TGTGTATCTGCATCAGCGGTTCTCTAGG 964

RESULT 6

AAT96713

ID AAT96713 standard; DNA; 2117 BP.

XX AAT96713;
AC
XX
DT 22-APR-1998 (first entry)
XX
DE Human hyaluronate synthetase coding sequence.
XX
KW Hyaluronate synthetase; human; hyaluronic acid; drug preparation;
KW cosmetic preparation; gene therapy; carcinogenesis; ss.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 149..1780
FT /*tag= a
XX
XX
PN WO9738113-A1.
XX
PD 16-OCT-1997.
XX
XX 31-MAR-1997; 97WO-JP001111.
XX
XX 05-APR-1996; 96JP-00084326.
PR 30-APR-1996; 96JP-00109663.
XX
XX (SEBK) SEIKAGAKU CORP.
PA
XX Itano N, Kimata K;
PI
XX
XX WPI; 1997-512726/47.
DR P-PSDB; AAW36503.
XX
XX DNA encoding human hyaluronate synthetase - for industrial scale
PT production of hyaluronic acid used in generating anti-carcinogenic drugs
PT or for cosmetics.
XX
XX Claim 3; Page 23-27; 35pp; Japanese.
XX
XX This sequence encodes a human hyaluronate synthetase, and is the coding
CC sequence of the invention. The encoded enzyme is useful for industrial
CC scale production of hyaluronic acid for use in the preparation of drugs
CC and cosmetics. The drugs can also be used in compositions for the
CC treatment of disorders involving the lowering of hyaluronic acid
CC production. The peptides may be used for the preparation of antibodies
CC recognising hyaluronate synthetase, e.g. for diagnostic purposes.
CC Antisense DNA or RNA corresponding to the DNA can be used in gene therapy
CC treatment of carcinogenesis
XX
XX Sequence 2117 BP; 306 A; 673 C; 703 G; 435 T; 0 U; 0 Other;
SQ

Query Match 84.5%; Score 905.4; DB 2; Length 2117;
Best Local Similarity 98.8%; Pred. No. 1.1e-148;
Matches 912; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 7 CAGCAGGACGGCCCAAGCCCACTCTCGAGCCGCGCTGCTCGGGCTGGCCGGAGG 66
Db 50 CAGCAGGACGGCCCAAGCCCACTCTCGAGCCGCGCTGCTCGGGCTGGCCGGAGG 109
Qy 67 GTGCTGACCAATCGGCTTTCGCGCTGCTCATCTCTGGGCTCATGACCTGGGCTACGCCGCC 126
Db 110 GTGCTGACCAATCGGCTTTCGCGCTGCTCATCTCTGGGCTCATGACCTGGGCTACGCCGCC 169
Qy 127 GGGGTGGCGCTGGGCTTCGATCGCTACGGCTCTCTGGGCTTCGGGCTTCAGGGGCTTC 186
Db 170 GGGGTGGCGCTGGGCTTCGATCGCTACGGCTCTCTGGGCTTCGGGCTTCAGGGGCTTC 229
Qy 187 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTCGCTACCTGGAGCACCGCGGGTGGCG 246
Db 230 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTCGCTACCTGGAGCACCGCGGGTGGCG 289
Qy 247 GCGCGCGCGGGGGCGCTGGATGACGCCACCGCGCGAGTGTGGGCTGACCATCTCC 306
Db 290 GCGCGCGCGGGGGCGCTGGATGACGCCACCGCGCGAGTGTGGGCTGACCATCTCC 349

Qy 307 GCCTACGAGGAGACCCGCGGTACCTGGCGCAGTGCCCTGGCGTCCGCCGCCCTGCTG 366
Db 350 GCCTACGAGGAGACCCCGCTTACCTGGCGCAGTGCCCTGGCGTCCGCCGCCCTGCTG 409
Qy 367 TACCCGCGCGCGCGCTGCGCTCTCATGTGTGGATGGCAACCGCGCGAGACCTC 426
Db 410 TACCCGCGCGCGCGCTGCGCTCTCATGTGTGGATGGCAACCGCGCGAGACCTC 469
Qy 427 TACATGTCGACATGTTCCCGAGGCTTTCGCTGACGAGGACCCCGCGCTGCTG 486
Db 470 TACATGTCGACATGTTCCCGAGGCTTTCGCTGACGAGGACCCCGCGCTGCTG 529
Qy 487 GACGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGCGCGCGCGAGCC 546
Db 530 GACGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGCGCGCGAGCC 589
Qy 547 TATCGGAGGTGGAGGCGGAGGATCCCTGGGCGGCTGGCAGTGGAGGCGCTGCTGAGGACT 606
Db 590 TATCGGAGGTGGAGGCGGAGGATCCCTGGGCGGCTGGCAGTGGAGGCGCTGCTGAGGACT 649
Qy 607 CGCAGGTGCTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCATGTACAGCC 666
Db 650 CGCAGGTGCTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCATGTACAGCC 709
Qy 667 TTCAGGCGCTCGGAGATTCCGTGGGACTACGTGCGAGGTCTGTGACTCGACACAGGTTG 726
Db 710 TTCAGGCGCTCGGAGATTCCGTGGGACTACGTGCGAGGTCTGTGACTCGACACAGGTTG 769
Qy 727 GACCCCATGCGACTGCTGGAGCTCGTGGGCTACTGGACGAGGACCCCGGGTAGGGGCT 786
Db 770 GACCCCATGCGACTGCTGGAGCTCGTGGGCTACTGGACGAGGACCCCGGGTAGGGGCT 829
Qy 787 GTTGGTGGGACGTCGCGATCCCTTAACCTCTGAGCTCTCTGGGTCAGCTTCTTAAGCAGC 846
Db 830 GTTGGTGGGATGTCGCGATCCCTTAACCTCTGAGCTCTCTGGGTCAGCTTCTTAAGCAGC 889
Qy 847 CTCGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCACTGTGTA 906
Db 890 CTCGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCACTGTGTA 949
Qy 907 TCCTGCAATCAGCGGTTCTCTAGG 929
Db 950 TCCTGCAATCAGCGGTTCTCTAGG 972

RESULT 7
ADL13693/c
ID ADL13693 standard; DNA; 231222 BP.
XX
AC ADL13693;
XX
DT 06-MAY-2004 (first entry)
XX
DE Osteoarthritis-associated polymorphic nucleotide #225.
XX
KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
XX WO2003054166-A2.
XX
XX 03-JUL-2003.
XX
XX 19-DEC-2002; 2002WO-US041225.
XX
XX 20-DEC-2001; 2001US-0342603P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Jones KA, Schafer A;

CC a human eye cell. The present sequence encodes mouse HAS1 which is used
CC in the exemplification of the present invention
XX
SQ Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;

Query Match 63.3%; Score 678.2; DB 8; Length 1752;
Best Local Similarity 84.5%; Pred. No. 3.9e-109; Mismatches 128; Indels 18; Gaps 2;
Matches 793; Conservative 0;

Qy 10 CAGGACGCGCCCAAGCCCACTCCCTGCGAGCCCGCGCTGCTCGGCTGCGCCGAGGGTG 69
Db |||||
Qy 7 CAGGACATGCGCAAGCCCTCAGAGCGAGCGGTTGCTGCTGCTGCGCTGCGCCAGCGAGCA 66
Db |||||
Qy 70 CTGACCATCGCCTTCGCGCTGCTCATCTGCGGCTCATGACCTGGGCTAGCGCCGCGGG 129
Db |||||
Qy 67 CTCACGATCATCTTTGGCTGCTCATCTGCGGCTCATGACCTGGGCTAGCGCCGAGGC 126
Db |||||
Qy 130 GTGCGGCTGGGCTCGGCTAGCGCTCTGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 189
Db |||||
Qy 127 GTTCTCTGCGCTTCAGATCGCTATGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 186
Db |||||
Qy 190 TCAGCGACCTGCTGGGCGAGAGCTTCTGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 249
Db |||||
Qy 187 AGCGCACACCTAGTGGCGACAGAGCTTCTGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 246
Db |||||
Qy 250 GCGGCGCG-----GGGCGCGTGGATGAGCGCACCGCGCGGCTGGGCTGGGCG 294
Db |||||
Qy 247 GCTGCGGCGCGCTCTTTGGCGAAGGGCGGCTGGATGGGCGCTGCGGCTGGGCG 306
Db |||||
Qy 295 CTGACCATCTCCGCTTACCGAGGAGCGCGGCTGCTGCGCTTGGGCTTGGGCTTGGGCTTGGGCTT 354
Db |||||
Qy 307 CTCACCATCTGAGCTTACCAAGAGATCCCGCTTACCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 366
Db |||||
Qy 355 CGCGCCCTGCTGTACCGCGCGCGCTGCGGCTGCTGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 414
Db |||||
Qy 367 CGCGCCCTGCTGTACCGCGCACCGAGGTTACCGGCTGCTGCGCTTGGGCTTGGGCTTGGGCTTGGGCTT 426
Db |||||
Qy 415 GCGAGGACCTTACATGCTGACATGTTGCGGAGGTCTTGGCTGAGAGGAGCGCGCC 474
Db |||||
Qy 427 GCTGAGGATCTGATGCTGAGATGTTCCGAGAGGTCTTCCGCGGATGAGGAGCGCGCC 486
Db |||||
Qy 475 ACGTACGTGTGGGACGCACTACACAGCGCTTGGGAAACCGCGCGCG---CGGCGCG 531
Db |||||
Qy 487 ACTTATGTGGGATGGCACTACATCAGCCCTTGGGAAACAGCGGAGGCTACGGGCGCT 546
Db |||||
Qy 532 GTGGGCGCGGAGCTTATCGGAGGTGGAGCGGAGGATCTTGGGCGGCTGGCAGTGGAG 591
Db |||||
Qy 547 GTCGTTGAAGTGCTTACCGGAGGTGGAGCGGAGGACCCCGGCGGTTGGGCGTGGAG 606
Db |||||
Qy 592 GCGTGTGAGGACTCGGAGGTGCTGTGCTGCGGAGCGGCTGCGGCGGCGCAAGCGCGAG 651
Db |||||
Qy 607 GCGTGTGAGAAACAGCGAGGTGCTGTGCTGCTCAGCGTTGGGCGGCGAAACGTTGAG 666
Db |||||
Qy 652 GTCATGTACACAGCTTTCAGGCGCTCGGAGATTCGGTGGACTACGTGAGCTGTGAG 711
Db |||||
Qy 667 GTCATGTACAGCTTTCAGGCGACTGGCGGACTCCGTTGGACTACGTGAGGCTGTGAG 726
Db |||||
Qy 712 TCAGACACAAGGTTGGACCCCATGCTGCTGAGCTCTGTCGCGGTACTGGACGAGGAC 771
Db |||||
Qy 727 TCAGACACAAGACTAGACCCCATGCTGCTGAGCTTGTGCGAGTGTGTTGATGAAGAC 786
Db |||||
Qy 772 CCGCGGTAGGGCTGTGTTGGGAGAGTGGGATCTTAACTTGGACTCTGGGCTCCTGGTC 831
Db |||||
Qy 787 CCGCGGTAGGGCTGTGTTGGAGGGGATGTGAGGATCTTAACTTGGACTCTGGGCTCCTGGTC 846
Db |||||
Qy 832 AGCTTCTAGAGCGCTTGGGATCTGAGTGGCTTCAATGTGAGCGGCTTGTGAGAGC 891
Db |||||
Qy 847 AGCTTCTTGAAGCTCTTGGATCTGGGATGCTTCAATGTGGAACGAGCTTGTGAGAGC 906
Db |||||
Qy 892 TACTTCCACTGTGTPATCTTGCATCAGCGGTTCTTAGGT 930
Db |||||
Qy 907 TACTTCCACTGTGTPCTTGCATCAGTGGTCTCTGGGT 945
Db |||||

RESULT 9

AAD59442
ID AAD59442 standard; DNA; 1752 BP.

XX AC AAD59442;

XX DT 18-DEC-2003 (first entry)

XX Mouse hyaluronan synthase (HAS) 1 DNA.

XX Mouse; hyaluronic acid; HA; hyaluronan synthase; HAS; dry eye syndrome;
XX eye cell; osteoarthritis; gene therapy; gene; ds.

XX OS Mus sp.

XX Key Location/Qualifiers

XX CDS 1..1752

XX FT /*tag= a

XX FT /product= "Mouse hyaluronan synthase 1"

XX US2003087850-A1.

XX PN 08-MAY-2003.

XX PF 10-JUL-2001; 2001US-00902939.

XX PR 10-JUL-2001; 2001US-00902939.

XX (DEHA/) DEHAZYA P.

XX (CHEN/) CHEN W.

XX Dehazy P, Chen W;

XX WPI; 2003-755151/71.

XX P-PSDB; AAE39152.

XX Dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate

XX comprising derivatized hyaluronic acid cross linked to nucleic acid

XX encoding hyaluronan synthase useful for treating dry eye syndrome.

XX Claim 1; Page 12-13; 31pp; English.

XX The invention relates to dihydrazide derivatised hyaluronic acid (HA)/

XX nucleic acid bioconjugate comprising derivatised HA cross linked to

XX nucleic acid encoding hyaluronan synthase (HAS). The invention is useful

XX for treating an eye cell of an individual. It is useful for treating

XX dry eye syndrome and osteoarthritis of the particular joints. The

XX invention is also useful in gene therapy. The present sequence is mouse

XX HAS1 DNA

XX SQ Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;

XX Query Match 63.3%; Score 678.2; DB 10; Length 1752;

XX Best Local Similarity 84.5%; Pred. No. 3.9e-109;

XX Matches 793; Conservative 0; Mismatches 128; Indels 18; Gaps 2;

Qy 10 CAGGACGCGCCCAAGCCCACTCCCTGCGAGCCCGCGCTGCTCGGCTGCGCCGAGGGTG 69

Db |||||

Qy 7 CAGGACATGCGCAAGCCCTCAGAGCGAGCGGTTGCTGCTGCTGCGCTTGGGCTAGCGCGAGCA 66

Db |||||

```
QY 250 GCGGCGCG-----GGGGCGCTGGATGACAGCCACCGCGCGCTGGCGG 294
Db |||||
QY 247 GCTGCGCGCGCTCTCTTGGCGAAGGGCGCCCTGGATGCGGCCACTGACAGCGTGGCA 306
Db |||||
QY 295 CTGACCATCTCCGCTTACAGAGGACCCCGGTACCTGCGCGAGTGGCTGGCGTCCGCC 354
Db |||||
QY 307 CTCACCATCTCAGCTTACCAAGAGGATCCCGCTTACCTGCGCGAGTGTGACCTCCGCG 366
Db |||||
QY 355 CGCGCCCTCTGTACCCGCGCGCGCTGCGCGTCTCTCATGTGTGTGATGCCAACCCG 414
Db |||||
QY 367 CGCGCCCTCTGTACCCGCGCGAGGTACCGGTGCTCATGTGTGTGACCGCAACCG 426
Db |||||
QY 415 GCGGAGGACCTTACATGTGTGACATGTTCCGCGAGTCTTCTGCTGACGAGACCCCGC 474
Db |||||
QY 427 GCTGAGGATCTTACATGTGTGACATGTTCCGAGAAGTCTTCCCGATGAGGACCCCGC 486
Db |||||
QY 475 AGCTAGTGTGGAGCGCACTACACAGACCCCTGGAAACCCGCGCGG---CGGGCGG 531
Db |||||
QY 487 ACTTATGTGTGGATGGCACTACCATCAGCCCTTGGAAACCGAGCGAGCTACGGGCGCT 546
Db |||||
QY 532 GTGGGCGCGGAGCTTATCGGAGGTGAGGCGGAGGATCTTGGGCGGCTGGCAGTGGAG 591
Db |||||
QY 547 GTCGCTGAAGTGCCTTACCGGAGGTGAGGCGGAGACCCCGCGGCTTGGCGGTGGAG 606
Db |||||
QY 592 GCGCTGTGAGGACTCGCAGTGTGCTGTGCTGCGGCGACGCTGGGGCGGCAAGCGCGAG 651
Db |||||
QY 607 GCGCTGTGAGAACACGCGAGTGTGCTGTGCTGAGGCTTGGGGCGGCAACGCTGAG 666
Db |||||
QY 652 GTCATGTACACAGCTTCAAGCGCTCGAGATTCGGTGGACTACGTGCAAGTCTGTGAC 711
Db |||||
QY 667 GTCATGTACACAGCTTCAAGCGACTGGGCGACTCGTGGAGTACGTGCAAGTCTGTGAC 726
Db |||||
QY 712 TCGGACACAAAGTTGGACCCCATGCTGCTGAGCTGCTGCGGTACTTGGACGAGGAC 771
Db |||||
QY 727 TCAGACACAGACTAGACCCCATGCTGCTGAGCTGTGCGAGTGTGATGAGAC 786
Db |||||
QY 772 CCGCGGTAGGGCTGTGGTGGGACGFGCGGATCCTTAAACCTCTTGAATCTCTGGGTC 831
Db |||||
QY 787 CCGCGGTAGGGCTGTGGAGGGATGTGAGGATCTTAAACCTCTTGAATCTCTGGGTC 846
Db |||||
QY 832 AGCTTCTTAAAGCTGCGATCTGCTGAGTGTGCTTCAATGTGGAGCGGGCTTGTGAGC 891
Db |||||
QY 847 AGCTTCTTGAAGCAGTCTTGAATCTGCTGAGTGTGCTTCAATGTGGAGCGAGCTTGTGAGC 906
Db |||||
QY 892 TACTTCCACTGTATCTCTGATCAGCGGTCTCTTAGGT 930
Db |||||
QY 907 TACTTCCACTGTGTCTCTGATCAGTGTGCTCTGAGGT 945
Db |||||
```

RESULT 10

```
AAT91655
ID AAT91655 standard; DNA; 2102 BP.
XX
AC AAT91655;
XX
XX
DT 17-OCT-2003 (revised)
DT 19-DEC-1997 (first entry)
XX
DE Mouse hyaluronate synthase genomic DNA.
XX
KW Hyaluronate synthase; hyaluronic acid; polysaccharide; ds.
XX
OS Mus sp; (strain FM3A).
XX
FH Location/Qualifiers
FT 49..1800
FT /*tag= a
FT /product= "Hyaluronate synthase"
XX
PN JP09224674-A.
XX
XX 02-SEP-1997.
XX
```

```
PF 26-FEB-1996; 96JP-00038336.
XX
PR 26-FEB-1996; 96JP-00038336.
XX
PA (KAGG) KAGAKU GIJUTSUCHO CHOKAN KANBO.
XX
XX WPI; 1997-484102/45.
DR P-PSDB; AAW30704.
XX
XX Hyaluronate synthase isolated from mouse cells - useful for large-scale
XX production of hyaluronic acid.
XX
XX Claim 2; Page 10-13; 15pp; Japanese.
XX
XX A mouse-derived cell strain, FM3A P15-A, was cultured in Eagle minimal
XX medium containing 10 % heat-inactivated bovine serum, twice concentration
XX of amino acids and vitamins and penicillin and streptomycin at 37 degrees
XX Celsius. The culture was subjected to immobilised erythrocyte exclusion
XX to examine the extent of extracellular formation of hyaluronic acid
XX matrix. Cells which showed high formation were recovered and named FM3A
XX HAI. Cells synthesising no hyaluronic acid were recovered and named HAS-.
XX The HAS- cell in which polyoma large T antigen was expressed was prepared
XX (HAS- P cell ). Poly(A)+ RNA was isolated from FM3A HAI and cDNA was
XX prepared and was used for constructing a library in HAS- P cells. Cells
XX having hyaluronic acid synthetic activity were selected from the
XX transformants and plasmid DNA was recovered and amplified in E.coli. The
XX resulting genomic DNA sequence codes for hyaluronate synthase having a
XX sequence of 583 amino acids which is used for large-scale production of
XX hyaluronic acid, a component of drugs and cosmetics. (Updated on 17-OCT-
XX 2003 to standardise OS field)
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
```

Query Match 63.3%; Score 678.2; DB 2; Length 2102;

Best Local Similarity 84.5%; Fred. No. 3.9e-105;

Matches 793; Conservative 0; Mismatches 128; Indels 18; Gaps 2;

```
QY 10 CAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCTGCTCGGCTGCGCGGAGGGTG 69
Db |||||
QY 55 CAGGACATGCCAAGCCCTCAGAGCAGCGGTGCTGCTCTGGCTTGGCAGGCGAGCA 114
Db |||||
QY 70 CTGACCATCGCTTGGCCCTCTCATCTGGGCTCATGACCTGGGCTTACGCCCGCGG 129
Db |||||
QY 115 CTCAGATCATCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTACGCCAGGC 174
Db |||||
QY 130 GTGCGCTGGGCTCGGATCGCTACGGCTCTGGGCTTGGGCTTACGCCGCTTCTT 189
Db |||||
QY 175 GTTCTCTGGCTTCTGATCGCTATGAGTCTCTGGCTTGGGCTTATGGGCGCATTCCTC 234
Db |||||
QY 190 TCAGCGCACCTGGTGGCGCAGAGCCTCTTCCGCTACCTGGAGCACCGCGGTGGCGG 249
Db |||||
QY 235 AGCGCACACCTAGTGGCACAGAGCTCTTCCGCTTACCTTGGAGCACCGAGGCTGGCAGCG 294
Db |||||
QY 250 GCGGCGCG-----GGGGCGCTGGATGACAGCCACCGCGCGAGTGTGCGG 294
Db |||||
QY 295 CTGACCATCTCCGCTTACAGAGGACCCCGGTACCTTGGCGAGTGGCTTGGCGTCCGCC 354
Db |||||
QY 355 CTCACCATCTCAGCTTACCAAGAGGATCCCGCTTACCTGCGCGCAAGTGTGACCTCCGCG 414
Db |||||
QY 355 CGCGCCCTCTGTATCCCGCGCGGCTGCGCGTCTCTCATGTGTGTGATGGCAACCGC 414
Db |||||
QY 415 CGCGCTTGTGTATCCCGCACACGAGGTTCAGCGTGTCTCATGTGTGTGAGCGGCAACCGC 474
Db |||||
QY 415 GCGGAGGACCTCTATAGTGTGACATGTTCCGCGAGTGTCTTCTGCTGACGAGACCCCGC 474
Db |||||
QY 475 GCTGAGGATCTGTATGTTGGACATGTTCCGAGAAGTCTTTCGCGATGAGGACCCCGC 534
Db |||||
QY 475 ACGTACGTGTGGAGCGCAACTACCAAGCCCTGGGAAACCCGCGCGG---CGGGCGG 531
Db |||||
QY 535 ACTTATGTGTGGATGGCAACTACCATCAGCCCTGGGAAACCGAGCGGAGGCTACCGGCGCT 594
Db |||||
```


PV		W0200218632-A2.
XX		
PD		07-MAR-2002.
XX		
PF		01-SEP-2001; 2001WO-EP010074.
XX		
PR		01-SEP-2000; 2000DE-01043826.
XX		
PR		05-SEP-2000; 2000DE-01044543.
XX		
PA		(EPIG-) EPIGENOMICS AG.
XX		
PI	Olek A,	Piepenbrock C, Berlin K, Guetig D;
XX		
DR	WFI;	2002-371829/40.
XX		
PT	Determining the degree of cytosine methylation in genomic DNA, useful for	
PT	diagnosis and prognosis, comprises selective hybridization of amplicons	
PT	from chemically treated DNA.	
XX		
PS	Claim 12; 56pp + Sequence Listing; 56pp; German.	
XX		
CC	This invention describes a novel method for determining the degree of	
CC	methylation of a particular cytosine in a motif 5'-CpG-3', present in a	
CC	genomic sample of DNA. The sample is treated chemically to convert	
CC	cytosine (C) but not methylated C, to uracil, then part of the genomic	
CC	DNA that contains the target C is amplified to form a labeled amplicon.	
CC	The amplicon is hybridised to two classes, each with at least one member,	
CC	of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the	
CC	degree of hybridisation to both classes is determined from the label on	
CC	the amplicon. From the ratio of labels hybridised to the two classes of	
CC	oligomers, the degree of methylation is calculated. The method is used:	
CC	(i) for diagnosis and/or prognosis of side effects of therapeutic drugs	
CC	and of a wide range of diseases, e.g. cancer, disorders of the central	
CC	nervous, cardiovascular, gastrointestinal and respiratory systems etc.,	
CC	particularly by detecting mutations or single nucleotide polymorphisms	
CC	(SNPs); and (ii) for differentiation of cell or tissue types and for	
CC	investigating cell differentiation. The method allows the methylation	
CC	status of many C residues to be determined simultaneously. ABQ13410-	
CC	ABQ54121 represent genomic DNA sequences used to illustrate the method	
CC	for determining the degree of cytosine methylation described in the	
CC	disclosure of the invention	
XX		
SQ	Sequence 662 BP; 255 A; 247 C; 80 G; 80 T; 0 U; 0 Other;	
	Query Match 39.4%; Score 422; DB 6; Length 662;	
	Best Local Similarity 77.3%; Pred.No. 1.6e-64;	
	Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;	
Qy	34 GCAGCCCGCGCTGTCTCCGGCTTGCAGGGGTGCTGACCATCGCTTCGCCCTGCTC 93	
Db	662 GTAGTTGTGCTTGTTTCGGTTTTGGTTCGAGGGTGTGATTATCGTTTTCGTTTCTTT 603	
Qy	94 ATCCTGGGCGCTCATGACCTGGGCTTAGCGCGCGGGGTGCGCTGCGCTCCGATCGGTAC 153	
Db	602 ATTTCGGGTTTATGATTTGGGTTTACGTGCTCGGGGTGTGTTGGTTTTTCGATCGTTAC 543	
Qy	154 GGCGCTCTGSCCTTCGSCCTCTACGGGGCTTCTCTTTACGCGCACCTGTGTGGCGCAGAGC 213	
Db	542 GGTTTTTTGGTTTTTCGGTTTTTACGGGGTTTTTTTTTTTACGCTATTTTGGTGGCGTAGAGT 483	
Qy	214 CTCCTCCGCTACCTGGAGCACCGCGGGGTGGCGCGCGCGGGGGCGCGCTGGATGCA 273	
Db	482 TTITTTCCGGTATTTTGGAGTATCGCGGGTGGCGCGCGCGGGGTGCTTTGGATGA 423	
Qy	274 GCCACC CGCGCAGTGTGGCGCTGACA CATCTCCGCTTACAGAGGACACCCCGGTACTGT 333	
Db	422 GTTATCGCGGTAGTGTGGCGTTGATTAATTTTCGTTTATTTAGGAGGATTTTCGGGTATT 363	
Qy	334 CGCCAGTGCCTGCGCTCCGCGCGCCCTGCTGTA CC CGCGCGCGGGCTGCGCGCTCCT 393	
Db	362 CGTTAGTGTGTTGCGCTTCGTTTCGCTTTTGTGTA TTCGCGCGCGGGTTCGCGGTTTTT 303	
Qy	394 ATGGTGGTGATGGCAACCGCGCGCGGAGGACCTCTACATGCTCGACATGTTCCGCGAGGTC 453	

Qy	154	GGCTCTCTGCGCTTCGGCGCTCTACGGGGCTTCCTTTACGCGCACTGTGTGGCGCAGAC	213
Db	542	GGTTTTTTTGGTTTTTCGTTTTTACAGGGTTTTTTTTTTTACGCGTATTTGTGTGGCGTAGAGT	483
Qy	214	CTCTTCGCGTACCTGTGAGCACCGGCGGCTGTGGCGCGCGGGGGCGCTGGATGCA	273
Db	482	TTTTTTTCGCGTATTTTGAGATATCGCGGGGTGGCGCGCGCGGGGGTCTTTGGATGTA	423
Qy	274	GCACCGCGCGCACTGTGGCGCTGACCATCTCGCGCTACCCAGGAGGACCCCGCGGTACCTG	333
Db	422	GTTATCGCGCTAGTGTGCGTTGATTAATTTTCGTTTATTTAGGAGGATTTTCGCGTATTTG	363
Qy	334	CGCGAGTGCCTGTGCGTTCGGCGCGCGCGCTGTGTATCCGGCGCGCGGCTGCGCGGCTC	393
Db	362	CGTTAGTGTGTGCGTTCGTTTCGCGTTTGTGTATTCGCGCGCGCGGTTGCGCGTTTTT	303
Qy	394	ATGGTGGTGGATGGCAACCGCGCGGAGGACCTCTACATGCTCGACATGTTCCGCGAGGTC	453

302	ATGTTGGTGGATGTTAAATCGCGTCGAGGATTTTATATGGTCGATATGTTTCGCGAGGTT	243
Db		
Qy	TTTCGCTGACGAGGACCCCGCCACGTCAGTGTGGGACGGCAACTACACACAGCCCTGGGAA	513
Db		
242	TTTCGTTGACGAGGATTTTCGTTACGTCAGTGTGGGACGGTAATTATATTAGTTTGGGNA	183
Qy	CCCCGCGCGCGCGCGCGTGGCGCCGAGAGCCTATCCGGAGGTCGAGCGCGAGGATCCT	573
Db		
182	TTTCGCGCGCGCGCGCGTGGCGCTCGGAGTTTATCCGAGGTCGAGCGCGAGGATTTT	123
Qy	GGGCGGCTGGCAGTCGAGGGCGCTGCTGAGGACTCCGACGTCGCTGCGTGGCGCGAGCGC	633
Db		
122	GGGCGGTTGGTAGTGAGGCGGCTGGTGAGGATTCGTAGTGCCTGTCGCTGGCGCGTAGCGT	63
Qy	TGGGCGCGCAAGCGCGAGGTCATGTACACAGCCTTCAAGCGCCTCCGAGAGATTCGGTGGAC	693
Db		
62	TGGGCGGTAAAGCGGAGGTTATGTATAGTTTATTAGGCGCTTCGGAGATTCGGTGGAT	3
Qy	694 TA 695	
Db	2 TA 1	

Search completed: March 13, 2005, 03:46:24
Job time : 702.034 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 02:13:47 ; Search time 5058.4 Seconds
(without alignments)
10259.279 Million cell updates/sec

Title: US-10-672-399-3
Perfect score: 1071
Sequence: 1 atgacagcagcagcgcgc.....ccggcaccatgctggatga 1071

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	927.4	86.6	2088	9 HSU59269	US9269 Human hyalu
2	927.4	86.6	2116	6 AR220003	AR220003 Sequence
3	915.2	85.5	2119	9 BC035837	BC035837 Homo sapi
4	915	85.4	2108	9 HUMHAS	D84424 Homo sapien
5	915	85.4	2117	6 AR137038	AR137038 Sequence
6	854.4	79.8	2109	9 AY463695	AY463695 Papio anu
7	697.2	65.1	229155	9 AC018755	AC018755 Homo sapi
8	679.8	63.5	2095	10 AB097568	AB097568 Rattus no
9	678.2	63.3	2102	6 E13681	E13681 DNA encodin
10	678.2	63.3	2102	6 E28454	E28454 Hyaluronate
11	678.2	63.3	2102	6 E30971	E30971 Hyaluronate
12	678.2	63.3	2102	6 E34326	E34326 DNA for gen
13	678.2	63.3	2102	10 MUSHAS	D82964 Mus musculu
14	669	62.5	193986	2 AC130783	AC130783 Pan trogl
15	635.4	59.3	185623	2 AC137058	AC137058 Papio anu
16	614.8	57.4	1581	6 CQ731079	CQ731079 Sequence
17	492	45.9	293184	2 AC079487	AC079487 Mus muscu
18	490.8	45.8	72955	2 AC108651	Continuation (4 of
19	490.8	45.8	257325	2 AC116203	AC116203 Rattus no

c	20	369.2	34.5	611	11	BV099313	BV099313 RPAMMSQ0
c	21	368.2	34.4	611	11	BV160501	BV160501 RPAMMSQ0
	22	302.4	28.2	1375	5	AY437407	AY437407 Danio rer
	23	290.6	27.1	1659	4	AB055979	AB055979 Orcystolag
	24	274.2	25.6	550	11	BV161377	BV161377 RPAMMSQ0
	25	272.8	25.5	1662	6	CQ723590	CQ723590 Sequence
	26	272.8	25.5	1994	9	AF234839	AF234839 Homo sapi
	27	272.8	25.5	4049	9	AF232772	AF232772 Homo sapi
	28	272.8	25.5	4220	6	CQ776471	CQ776471 Sequence
	29	265.4	24.8	3995	4	AB159675	AB159675 Sub scrof
	30	250	23.3	2093	5	BC047963	BC047963 Xenopus l
	31	248.4	23.2	2950	5	XELDG424	M22249 X.laevis DG
	32	243.2	22.7	2465	10	AB097569	AB097569 Rattus no
	33	240.4	22.4	2234	5	AF015780	AF015780 Xenopus l
	34	238.8	22.3	2245	5	BC077983	BC077983 Xenopus l
	35	236.2	22.1	1665	6	E30973	E30973 Hvaluronic
	36	236.2	22.1	1665	6	AR373269	AR373269 Sequence
	37	236.2	22.1	5919	6	CQ777380	CQ777380 Sequence
	38	236.2	22.1	5919	10	MMU86408	U86408 Mus musculu
	39	207.4	19.4	1229	6	CQ776472	CQ776472 Sequence
	40	207.4	19.4	1229	9	BC021853	BC021853 Homo sapi
	41	195.6	18.3	605	11	BV161384	BV161384 RPAMMSQ0
c	42	194.4	18.2	215158	5	BX539313	BX539313 Zebrafish
c	43	193	18.0	603	11	BV161386	BV161386 RPAMMSQ0
c	44	193	18.0	621	11	BV160507	BV160507 RPAMMSQ0
c	45	192.4	18.0	621	11	BV099322	BV099322 RPAMMSQ0

ALIGNMENTS

RESULT 1
HSU59269
LOCUS HSU59269 2088 bp mRNA linear PRI 24-SEP-1996
DEFINITION Human hyaluronan synthase mRNA, complete cds.
ACCESSION U59269
VERSION U59269.1 GI:1556464
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2088)
AUTHORS Shyjan,A.M., Heidin,P., Butcher,E.C., Yoshino,T. and Briskin,M.J.
TITLE Functional cloning of the cDNA for a human hyaluronan synthase
J. Biol. Chem. 271 (38), 23395-23399 (1996)
MEDLINE 96394438
PUBMED 8798544
REFERENCE 2 (bases 1 to 2088)
AUTHORS Briskin,M.J. and Shyjan,A.M.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-1996) LeukoSite Inc., 215 First Street,
Cambridge, MA 02142, USA

FEATURES

source
1..2088
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
36..1772
/codon_start=1
/product="hyaluronan synthase"
/protein_id="AAC50706.1"
/db_xref="GI:1556465"

CDS

/translation="MRQDAKPTPAARRCSGLARRVLTAFALLILGLMTWVAAGV
PLASDRYGLAFLGFLGAPLVAHLVAQSLFAYLEHRRVAAAARGPDRATASVALTIS
AYQEDPAYLRQCLASAPALYPRARLVLMVDGNRAEDLYMDFMREVFADDPATY
VWDGNYPQPEPAAGAGAGAVEAEDEGLRVALEALVTRCVCVQAWGGKREY
MYTAFKALGSDVDYQVCDSDTRDPMALLELVLDPRVAGVGDVRLNPLDSW
VSFLSSLYVWAFNVAFRCOSYFHCVSCI SPGLGLYRNLLIQGLEAWYKQFLGTHC
TFGDDRLHTRMLSMGYATKYTSRSRCYSETPSSFLWLSQTRWSKSYFREWLYNAL
WMMKHAWMTTEAVVSGLFFVFFAVLRLFYAGRPWALLLVLLVLCVQVLAKAFAA
WLKGLRMVLLSLYAPLYMCLGLPAKFLALVTMNSQMGTSGRKRLAANYVPLPLAL
WALLLLGLVRSVAHEARDWSGSRAAEAYHLAAGAGYVGVVWVAMLTLYVWGVRL

Db 756 AGTTGGACCCCATGCGACTGCTGGAGTCTGTCGGGTACTGGACGAGACCCCGGTA 815
 Qy 781 GGGGCTGTTGGTGGGACGTGGGATCCTTAACCTCTGGACTCTGGTCTGCTTCTCTA 840
 Db 816 GGGGCTGTTGGTGGGACGTGGGATCCTTAACCTCTGGACTCTGGTCTGCTTCTCTA 875
 Qy 841 AGCAGCTGCATATCTGGGTAGCTTCAATGTGGAGCGGGTGTGACAGCTTACTTCCAC 900
 Db 876 AGCAGCTGCATATCTGGGTAGCTTCAATGTGGAGCGGGTGTGACAGCTTACTTCCAC 935
 Qy 901 TGTGTATCTTCATCAGCGGTCTCTAGG 929
 Db 936 TGTGTATCTTCATCAGCGGTCTCTAGG 964

RESULT 3
 BC035837
 LOCUS
 DEFINITION Homo sapiens hyaluronan synthase 1, mRNA (cDNA clone MGC:46218
 IMAGE:5589083), complete cds.
 ACCESSION BC035837
 VERSION
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2119)
 Strausberg, R.
 Direct Submission
 Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nigr.nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAP Plate: 79 Row: f Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504338.

FEATURES

Location/Qualifiers
 1..2119
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:46218 IMAGE:5589083"
 /tissue_type="Ovary, pooled from 3 adults"
 /clone_lib="NIH MGC_125"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..2119
 /gene="HAS1"
 /note="synonym: HAS"
 /db_xref="LocusID:3036"
 /db_xref="MIM:601463"
 42..1775
 /gene="HAS1"
 /codon_start=1
 /product="hyaluronan synthase 1"
 /protein_id="AAH35837.1"
 /db_xref="GI:23243102"
 /db_xref="LocusID:3036"
 /db_xref="MIM:601463"
 /translation="MRQDAKPTPAARRCSGLARRVLTAFALLILGLMTWAYAAGVP
 LASPYGLAFGLYGLAFLSAHLVAQSLFAYLEHRRVAAARGPDLDAATRSVALTISA
 YQEPAYLRQCLASARALLVPRLRLVLMVVDGNRAEDLYWMDRFREVPADPATYV
 WDGNYPWEPMAAAGAGAGAYREVAEDPGRLAVALRTRRCVCAQAGGKREVM
 RTYFKALGSDDYQVQCDSTRDLPMALLELRVLDEDPVAVGGDVRIINPLDSWV
 SFGLSLRYNFMNSMGYATKYTSRSRCYSETSPSLRWLSQOTRKSQVFRMLYNALW
 WRRHAWMYEAVVSGLFPFPAATVLRLEFYAGRPWALLWLLCVQVALAKAAPAAW
 LRGLRWLVSLSYAPLYMGLLPAKFLALVTMNSGWTSGRRKLAAHYVPLPLALW
 ALLLLGLVRSVAHEARADWSGPRAAEAHYHLAAGAGAYGVYVWAMLTLYWVGVRRLC
 RRRTGGYRVQV"

ORIGIN

Query Match 85.5%; Score 915.2; DB 9; Length 2119;
 Best Local Similarity 99.7%; Pred. No. 4.5e-112;
 Matches 917; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 10 CAGGACGCGCCCAAGCCACTCTCTCGAGCCCGCGCTGCTCGGCTCGGCCGGAGGCTG 69
 Db 48 CAGGACGCGCCCAAGCCACTCTCTCGAGCCCGCGCTGCTCGGCTCGGCCGGAGGCTG 107
 Qy 70 CTGACCATCGCTTCGCGCTGCTCATCTGCGGCTCATGACCTGGGCTACGCGCCGG 129
 Db 108 CTGACCATCGCTTCGCGCTGCTCATCTGCGGCTCATGACCTGGGCTACGCGCCGG 167
 Qy 130 GTGCGCTGGCTCGGATCGCTAGCGCTCTCGGCTTCGGCTTCGCGCTTCGCTTCCTT 189
 Db 168 GTGCGCTGGCTTCGGATCGCTAGCGCTCTCGGCTTCGGCTTCGCGCTTCGCTTCCTT 227
 Qy 190 TCAGCGACCTGTTGGCGCAGAGCTCTTTCGCTACCTTGGAGCACCGCGCGGTGGCGCG 249
 Db 228 TCAGCGACCTGTTGGCGCAGAGCTCTTTCGCTACCTTGGAGCACCGCGCGGTGGCGCG 287
 Qy 250 GCGCGCGGGGGCGCTGGATGACGCCACCGCGCGCAGGTGTGGCGCTGACCATCTCCGCC 309
 Db 288 GCGCGCGGGGGCGCTGGATGACGCCACCGCGCGCAGGTGTGGCGCTGACCATCTCCGCC 347
 Qy 310 TACGAGGAGACCGCGGTACTCGGCGCAGTGCCTGGGCTCGCGCGCGCTGCTGTGTAC 369
 Db 348 TACGAGGAGACCGCGGTACTCGGCGCAGTGCCTGGGCTCGCGCGCGCTGCTGTGTAC 407


```
/db_xref="GI:9454517"
/translation="MPADVNLSQRPQVLGPEKQDGSCEASVSFEDVTVDPSREBWOQL
DPAQCLRYDVMLELYSHLFAVGTHIPNPEVIFMLKEKEPRVEAEVSHQRCEREF
GLEIPQKEITSKCAQFQMDGVEFTRDGSWCSILBELRLDADRTRKKQONQIOPMSHSA
PFNKKTLENSCNSQKOPKMRTRPHLASSOKOPKCCLETESI.KLNLFVNGONESN
DTEQLDDVVGSGQLFSHSSDACSKNHTGTFTFCGNOCRAVCGHKQSLKQHQIHTQK
KPCDSCGGSFTRQKSHLFAQORLHSVGNLHCCGCGKAFQMLSVLFRDHTGDIIP
CICKECGKVFTRBELTHQKTRKPKYKCHDCGKAFQMLSLFRHQRTHSRKLYE
CECGKGFQMSNTLIIHQKHTGERQVACSECGKAFQKSTLSLHQRTHSQKSVCCI
ECGQAFIQKHALIWHQSHTGKPYQCHNCGKSFISKOLDIHHRIHTGKPYCSCDC
KKTOKSHLNIHOKHTGERHHVCECGKAFNOKSILSMHORLHTGKPYKCECGK
AFTSKSQFKEHORLHTGKPYVTECGKAFNRSNPHKHQIHTHRRFPVYCKGKAF
VQKSELTHQRIHMERKPEYECDCGKSFSAKQPLKQRIHTGERPVCSECGKAFNN
RSNENKQHTTHTRDKSKYSVKGFQK"
1097..1198
/rpt_family="MER99"
repeat_region
1559..1829
/rpt_family="AluJb"
repeat_region
complement(2130..2355)
/rpt_family="LIME"
repeat_region
complement(2359..2627)
/rpt_family="AluY"
repeat_region
complement(2816..3183)
/rpt_family="HAL1"
repeat_region
complement(3290..3598)
/rpt_family="AluSg"
repeat_region
complement(3611..3760)
/rpt_family="MIR"
repeat_region
complement(3766..4067)
/rpt_family="HAL1"
repeat_region
4098..4215
/rpt_family="AluJb"
repeat_region
complement(4339..4406)
/rpt_family="HAL1"
repeat_region
complement(4518..4724)
/rpt_family="HAL1"
repeat_region
complement(4803..5110)
/rpt_family="AluY"
repeat_region
5266..5561
/rpt_family="AluSg"
repeat_region
5635..5932
/rpt_family="AluJo"
repeat_region
5984..6120
/rpt_family="AluJb"
repeat_region
6121..6423
/rpt_family="AluSx"
repeat_region
6424..6579
/rpt_family="AluJb"
repeat_region
6647..6913
/rpt_family="LIME"
repeat_region
6936..7295
/rpt_family="MER33"
repeat_region
7700..7785
/rpt_family="L2"
repeat_region
complement(9216..9434)
/rpt_family="LIME"
repeat_region
complement(9435..9745)
/rpt_family="AluY"
repeat_region
complement(9746..9815)
/rpt_family="LIME"
repeat_region
complement(10009..10303)
/rpt_family="AluSp"
repeat_region
11783..12637
/rpt_family="L1MB2"
repeat_region
12180..12330
/gene="ZNF175"
repeat_region
13436..13619
/rpt_family="MIR"
repeat_region
16068..16112
/rpt_family="(TA)n"
repeat_region
complement(16113..16329)
/rpt_family="MER46A"
16330..16390
/rpt_family="(TA)n"
repeat_region
16403..16439
/rpt_family="(TATG)n"
repeat_region
16793..17013
/rpt_family="MER58A"
repeat_region
17024..17052
/rpt_family="(CA)n"
repeat_region
17311..17545
/rpt_family="MIR"
complement(18683..18896)
/rpt_family="MIR"
repeat_region
complement(18915..19183)
/rpt_family="AluY"
repeat_region
19336..19599
/rpt_family="MIR"
repeat_region
19811..19928
/rpt_family="GA-rich"
repeat_region
22181..23081
/rpt_family="L2"
repeat_region
complement(23692..23901)
/rpt_family="L2"
repeat_region
complement(23902..24201)
/rpt_family="AluSg1"
repeat_region
complement(24202..24651)
/rpt_family="L2"
repeat_region
26136..26434
/rpt_family="LTR12"
repeat_region
26435..26525
/rpt_family="LTR30"
repeat_region
26526..26643
/rpt_family="LTR12"
repeat_region
complement(26686..26884)
/rpt_family="AluSg/x"
repeat_region
27065..27240
/rpt_family="MIR"
repeat_region
complement(27238..27554)
/rpt_family="L2"
repeat_region
28406..28433
/rpt_family="GC-rich"
repeat_region
28641..28678
/rpt_family="GC-rich"
repeat_region
complement(28997..29259)
/rpt_family="MSR1"
repeat_region
complement(30219..30489)
/rpt_family="L2"
repeat_region
30753..30882
/rpt_family="FLAM_C"
repeat_region
31138..31416
/rpt_family="AluY"
repeat_region
complement(31547..31716)
/rpt_family="MER3"
repeat_region
32041..32112
/rpt_family="LIMEc"
repeat_region
32117..32516
/rpt_family="MLT2E"
repeat_region
35056..35118
/rpt_family="MIR"
repeat_region
complement(35578..35636)
/rpt_family="MER5A"
repeat_region
complement(35690..35820)
/rpt_family="MIR"
repeat_region
35872..35902
/rpt_family="(TG)n"
repeat_region
complement(35906..36209)
/rpt_family="AluJo"
65.1%; Score 697.2; DB 9; Length 229155;
Query Match
Best Local Similarity 99.6%; Pred. No. 8.6e-84;
Matches 699; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

Qy

7 CAGCAGACGCGCCCAAGCCCACTCTCGACCGCGCTCTCGGCTGCGCGAGG 66
|||||

Db	147050	CAGCAGGACGCGCCCAAGCCACTCTCTGACGCTGCGCGCTTGGCCCGGAGG	146991
Qy	67	GTGCTGACCATCGCCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCGGCC	126
Db	146990	GTGCTGACCATCGCCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCGGCC	146931
Qy	127	GGGGTGGCGCTGGCCTCCGATCGCTACGGCTCTCTGGGCTTCTGGGCTCTACGGGGCTTTC	186
Db	146930	GGGGTGGCGCTGGCCTCCGATCGCTACGGCTCTCTGGGCTTCTGGGCTCTACGGGGCTTTC	146871
Qy	187	CTTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTTCGCGTACCTTGGAGCACCGCGGGTGGCG	246
Db	146870	CTTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTTCGCGTACCTTGGAGCACCGCGGGTGGCG	146811
Qy	247	CGCGCGCGCGGGGGCGCGCTGGATGCGACACCGCGCGCAGTGTGGGCTGACCATCTCC	306
Db	146810	CGCGCGCGCGGGGGCGCGCTGGATGCGACACCGCGCGCAGTGTGGGCTGACCATCTCC	146751
Qy	307	GCCTACCAAGGAGGACCCCGCTACTCTGGCCAGTGCCTGGCGTCCGCGCGCCCTGCTG	366
Db	146750	GCCTACCAAGGAGGACCCCGCTACTCTGGCCAGTGCCTGGCGTCCGCGCGCCCTGCTG	146691
Qy	367	TACCCGCGCGCGGCTGCGGCTCTCTCATGTGTGGTGGATGGCAACCGCGCGAGGACCTC	426
Db	146690	TACCCGCGCGCGGCTGCGGCTCTCTCATGTGTGGTGGATGGCAACCGCGCGAGGACCTC	146631
Qy	427	TACATGTGTCACATGTTCCCGGAGGTTCTTCTGACGAGGACCCCGCCACGTACTGTGG	486
Db	146630	TACATGTGTCACATGTTCCCGGAGGTTCTTCTGACGAGGACCCCGCCACGTACTGTGG	146571
Qy	487	GACGGCAACTACCAACGAGCTTGGGAACCCCGCGCGCGCGGCTGGGCGCGCGAGCC	546
Db	146570	GACGGCAACTACCAACGAGCTTGGGAACCCCGCGCGCGCGGCTGGGCGCGCGAGCC	146511
Qy	547	TATCGGAGGTGGAGGCGGAGGATCTCTGGGCGCTTGGCAGTGGAGGCGCTGGTCAAGACT	606
Db	146510	TATCGGAGGTGGAGGCGGAGGATCTCTGGGCGCTTGGCAGTGGAGGCGCTGGTCAAGACT	146451
Qy	607	CGCAGGTGCTGTGCTGGCGCAGCGCTTGGGGCGGCGAAGCGCGAGGTCATGTACACAGCC	666
Db	146450	CGCAGGTGCTGTGCTGGCGCAGCGCTTGGGGCGGCGAAGCGCGAGGTCATGTACACAGCC	146391
Qy	667	TTCAGGCGCTCGAGATTCTGGTGGACTTACGTGACAGTCTCT 708	
Db	146390	TTCAGGCGCTCGAGATTCTGGTGGACTTACGTGACAGTCTCT 146349	
RESULT 8			
LOCUS	AB097568	2095 bp mRNA linear	ROD 28-APR-2004
DEFINITION	Rattus norvegicus HAS1 mRNA for hyaluronan synthase 1, complete cds.		
ACCESSION	AB097568		
VERSION	AB097568.1	GI:26453348	
KEYWORDS	Rattus norvegicus (Norway rat)		
SOURCE	Rattus norvegicus		
ORGANISM	Rattus norvegicus		
REFERENCE	1 Itano,N., Sawai,T., Ateumi,F., Miyaishi,O., Taniguchi,S., Kannagi,R., Hamaguchi,M. and Kimata,K. Selective expression and functional characteristics of three Mammalian hyaluronan synthases in oncogenic malignant transformation J. Biol. Chem. 279 (18), 18679-18687 (2004)		
	PUBMED 14724275		
	2 (bases 1 to 2095)		
REFERENCE	Itano,N.		
AUTHORS	Direct Submission		
TITLE	Submitted (05-DEC-2002) Naoki Itano, Aichi Medical University, Institute for Molecular Science of Medicine, Nagakute, Aichi, Aichi		
JOURNAL			

FEATURES		480-1195, Japan (E-mail: itano@amugw.aichi-med-u.ac.jp, Tel:81-52-264-4811 (ex.2095), Fax:81-561-63-3532)	
source		Location/Qualifiers	
		1..2095	
		/organism="Rattus norvegicus"	
		/mol_type="mRNA"	
		/db_xref="taxon:10116"	
gene		1..2095	
		/gene="HAS1"	
CDS		43..1794	
		/gene="HAS1"	
		/codon_start=1	
		/product="hyaluronan synthase 1"	
		/protein_id="BAC43730.1"	
		/db_xref="GI:26453349"	
		/translation="MRQMPKPEAARCCSLGARRLVITITFALLILGLMTWAYAGVP LADSPYGLAFLGYGLFSLARLYPRTRLRVLMVDGRADLDLYMDRPREFADED LITSAYQEDPTLYRQCLTSARALLYPRTRLRVLMVDGRADLDLYMDRPREFADED PATYVMDGNTHQWPEAEAGAVEGAYREVEDPGLRAVEALVRTRRCVCVAQRWG KREVMYATFALGDSDVYQVCDSTRLDPMALLELVRLVDLPRVGNVAGVDYRLN PLDSWVFLSLRYWAFNVERACQSYFHCVCSIGSGPLGILYRNNLLQGLFLEWYNQKF LGTHCTFGDDRHILNTRMLSMGYATKYTSRCSYSETSSFLRWLSQQTRSKSVFREK LYNALMHRHAWMTYEAVYGLFPFFVAATVLRLFYAGRPWALLWLLCVQGVALLK APTAALWRLGCLRMVLLSLYAPLYMCGLLPAKFLALVTMQSGMGTSGRKLAAAYPV LPALWALLLGLLRSVAQEVRAADNSGPSRAAEAYHLAAGASAYVAYVWIMLTIIYV GVRRLCRRRSGGVRVQV"	
ORIGIN		Query Match 63.5%; Score 679.8; DB 10; Length 2095;	
		Best Local Similarity 84.6%; Pred. No. 6.4e-81;	
		Matches 794; Conservative 0; Mismatches 127; Indels 18; Gaps 2;	
Qy	10	CAGACGCGCCCAAGCCACTCTCTGACGCGCGCGCTGCTCGGCGCTTGGCCCGGAGGGTG	69
Db	49	CAGGACATGCGCAAGCGCTCAGAGGCGAGCAGTGTGCTCGCGCTCTGGCCAGCGGGTG	108
Qy	70	CTGACCATCGCTTCCGCCCTCATCTCGGGCTCATGCTGGGCTTACGCGCGCGGG	129
Db	109	CTACGATACCTTTCGCCCTGCTATCTGGGCTCATGACCTGGGCTTACGCGAGGGA	168
Qy	130	GTGCGCTGCGCTCCGATCGCTACGGCTCTCTGCGCTTCCGCTTACGCGGCTTCTT	189
Db	169	GTACCTCTGCTTCTGATCCTATGGCTCTCTGCGCTTGGGCTCTATGGGCGTTCCTC	228
Qy	190	TCAGCGCACTGTGGCGGAGAGCGCTTCTGCGTACCTGAGAGCACCGCGGGTGGCGCG	249
Db	229	AGTGCACACCTAGTGGCACAGAGCCTCTTCGCTTACCTGGAGCACCGAAGGCTGACCG	288
Qy	250	GGCGCGCG-----GGGCGCGCTGGATGCGACCGCACCGCGCGCAGTGTGGCG	294
Db	289	GCTGCGCGCGCGCTTTTTCGAAGGGACCCCTGGATCGGCGCACTGCGCGCAGCGTGGCA	348
Qy	295	CTGACCATCTCCGCTTACACGAGGAGAACCCCGCGTACCTGGCGCCAGTGTGGCTCGGCC	354
Db	349	CTACCATCTCAGCCTACACGAGGAGACCCACTTACCTGGCGCAGTGTGACCTCCGCG	408
Qy	355	CGCGCTCTGTGTACCGCGCGCGCGCTGCGGCTCTCTATGTGTGTGATGTGCAACCGC	414
Db	409	CGCGCTTGTGTATCCCGCGCAGCGCGCTGCGGCTTATGTGTGTAGACGGCAATCGC	468
Qy	415	GGCGAGGACCTCTACATGGTTCGCGAGGCTTTCGCTGACGAGGACCCCGCC	474
Db	469	GCGGAGGATCTGTACATGTGTGGACATGTTCCGAGAGTCTTTCGAGATGAGACCTTGCC	528
Qy	475	ACGTACGTGTGGGACGGCAACTACCAACCGCTTGGGAACCCCGC---GGCGCGCGGGCG	531
Db	529	ACTATGTGTGGGATGGCAACTACCACTTGGGAGCCAGCTGAGGCGCGGGTGTCT	588
Qy	532	GTGGCGCGCGGAGCGCTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAG	591
Db	589	GTGGGTGAAGGTGCTTACCGCGAGGTGGAGGCTGAGGACCTTGGCGCGCTGGCGGTAGAG	648
Qy	592	GCCTGTGTAGGAGCTCGCAGGTGCTGTGCTGCGCGCAGCGCTGGGGCGCGCAAGCGCGAG	651

Db 649 GCGCTGGTGAGACCGCAGGTCGCTGCGTGGCTCAGCGTGGGTGGCAAGCGCGAA 708
Qy 652 GTCTGTATACACAGCCCTTCAAGCGCTCGAGAGATTGGTGGACTACTGTGAGGTCTGTGAC 711
Db 709 GTCTGTATACACCGCTTTCAGGCACTGGGTGACTCGTGGACTACTGTGAGGTCTGTGAC 768
Qy 712 TCGGACACAAGTTGGACCCCATGACACTGTGGAGCTCGTGGGTAATCTGGACGAGGAC 771
Db 769 TCAGACACAAGGTTAGACCCCATGACACTGTGGAGCTTGTGGAGTGTGATGAAGAC 828
Qy 772 CCCCGGTAGGGCTGTTGGTGGGACGTCGGATCCCTTAACCTCTGGACTCCCTGGGTC 831
Db 829 CCCCGGTAGGGCTGTTGGAGGAGATGAGAAATCCCTTAACCTCTGGACTCTGGGTC 888
Qy 832 AGCTTCTTAAGCAGGCTCGATCTACTGGGTAGGCTTCAATGTGGAGCGGCTTGTGAGAGC 891
Db 889 AGCTTCTTTCAGAGCGCTTCGATCTGGGTAGGCTTCAATGTGGAGCGGCTTGTGAGAGC 948
Qy 892 TACTTCCACTGTGTATCTTCGATCAGCGGTTCTCTAGGT 930
Db 949 TACTTTCATGTGTCTCTGATCAGTGTGCTCTGGGT 987

RESULT 9

LOCUS E13681 2102 bp DNA linear PAT 27-APR-1998
DEFINITION DNA encoding novel mouse hyaluronate synthetase.

ACCESSION E13681

VERSION E13681.1 GI:3252450

KEYWORDS JP 1997224674-A/1.

SOURCE Mus musculus

ORGANISM Mus musculus (house mouse)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2102)

AUTHORS Itano,N. and Kimata,H.

TITLE POLYPEPTIDE OF NEW HYALURONIC ACID-SYNTHETIC ENZYME AND DNA CODING

JOURNAL THE SAME

PATENT: JP 1997224674-A 1 02-SEP-1997;

SCIENCE & TECH AGENCY

OS Mus musculus (mouse)

PN JP 1997224674-A/1

PD 02-SEP-1997

PF 26-FEB-1996 JP 1996038336

PI ITANO NAKI, KIMATA HIROHARU

PC C12N15/09,C07H21/04,C07K14/47,C12N9/00//A61K48/00,C12N1/21,PC

(C12N9/00,

PC C12R1:91),(C12N9/00,C12R1:19),(C12N1/21,C12R1:19); CC

strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

FH Key Location/Qualifiers

FH source 1..2102

FT /organism='Mus musculus'

FT /cell_line='FM3A'

FT 5'UTR 1..48

FT CDS 49..1800

FT 3'UTR /product='Novel mouse hyaluronate synthetase'

FT 1801..2102.

FEATURES

source

1..2102

/organism='Mus musculus'

/mol_type='genomic DNA'

/db_xref='taxon:10090'

ORIGIN

Query Match 63.3%; Score 678.2; DB 6; Length 2102;
Best Local Similarity 84.5%; Pred. No. 1e-80;
Matches 793; Conservative 0; Mismatches 128; Indels 18; Gaps 2;

Qy 10 CAGGACGGCCCAAGCCCACTCTCTCAGACCCCGCCCGCTCTCGGCTGCCCCCGAGGGTG 69
Db 55 CAGGACATGCCAAAGCCCTCAGAGGCAGCGCTTCTCTCTGCGCTCGCCAGGGGAGCA 114
Qy 70 CTGACCATCGCTTCGCGCTGCTCATCTTGGGCTCATGACCTGGGCTTACGCCCGCGG 129
Db 115 CTCAGATCATCTTTTGGCCCTGCTCATCTTGGGCTCATGACCTGGGCTTACGCCCGG 174
Qy 130 GTGCGGCTGGGCTCCGATCGCTTACCGGCTTCTCGGCTTCTCGGCTTCTACGGGGCTT 189
Db 175 GTTCTCTTGGCTTCAGATCGCTATGGAATCTCTGGCTTCTGGCTCTATGCGGCTTCT 234
Qy 190 TCAGCGCACTGTGTGGCGAGAGCTTCTGGCTTCTGGCTTCTGGCTTCTGGCTTCTGG 249
Db 235 AGCGCACACCTAGTGGCACAGAGCTTCTGGCTTCTGGCTTCTGGCTTCTGGCTTCTGG 294
Qy 250 GCGGCGGCG-----GGGCGGCTGGATGACGACCGCGGCGAGTGTGGCG 294
Db 295 GCTGCGCGGCGCTCTTGGCGAAGGGGCGGCTTGGATGGGCGACCTGACGCGAGCGTGGCA 354
Qy 295 GTGACCATCTCGGCTTACGAGGAGACCGCGCTTACCTGCGGCTGCTGCGGCTGCGGCT 354
Db 355 CTCACCATCTCAGCTTACCAAGAGGATCCCGCTTACCTGCGGCTGCTTGGCTTCTGG 414
Qy 355 GCGGCTGCTGTGTACCGCGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
Db 415 GCGGCTGCTGTGTACCGCGCACACGAGGTTACGCGTGTCTCATGCTGCTGCGGCTGCG 474
Qy 415 GCGGAGGACCTTACATGCTGACATGTTCCGCGAGGCTTCTGCTGACGAGGACCGCGG 474
Db 475 GCTGAGGATCTGTATCATGCTGAGCATGTTCCGAGAGTCTTTCGCGGATGAGGACCG 534
Qy 475 ACCTGCTGTGGGCGGCAACTTACCACAGCCCTCGGAAACCGCGGCGGCGG---CGG 531
Db 535 ACTTATGTGTGGGATGGCACTTACCATCAGCCCTGGGAAACCGCGGAGGCTACGGGCG 594
Qy 532 GTGGCGCGGAGGCTATCGGAGGTTGAGGCGGAGGATCTTGGCGGCTGCGGCTGCGGAG 591
Db 595 GTGGGTGAAGGTGCTTACCGGAGGTGAGGCGGAGGACCGCGGCGGCTTGGCGGTGG 654
Qy 592 GCGCTGTTGAGGACTCGCAGGTGCTGCTGCGGCGAGCGCTGCGGCGGCGGCGGCGGAG 651
Db 655 GCGCTGTTGAGAAACACGCGAGGTGCTGCTGCGGCTGCGGCTGCGGCGGCGGCGGAG 714
Qy 652 GTCTGTATACACAGCTTCAAGCGCTCGGAGATTCCGTTGGACTAGCTGCGAGGTCTGT 711
Db 715 GTCTGTATACACAGCTTTCAGGCACTGGGCGAGCTCCGTTGGACTACTGTCAGGTT 774
Qy 712 TCGGACACAAGTTGGACCCCATGGSACTGTGGAGCTCGTGGGCTTACTGGACGAGGAC 771
Db 775 TCAGACACAAGACTAGACCCCATGCGACTGCTGGAGCTTGTGCGAGTGTGGATGAAGAC 834
Qy 772 CCCCGGTAGGGCTGTTGGTGGGACCGTGGGATCTTTAACCTCTGAGCTCTGGGTC 831
Db 835 CCCCGGTAGGGCTGTTGGAGGGGATGTGAGGATCTTTAACCTCTGAGCTCTGGGTC 894
Qy 832 AGCTTCTTAAGCAGCTCGGATCTGCGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGC 891
Db 895 AGCTTCTTGGAGCGTCTTCGATCTGCGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGC 954
Qy 892 TACTTCCACTGTGTATCTTCGATCAGCGGTTCTCTAGGT 930
Db 955 TACTTCCACTGTGTCTCTGATCAGTGTGCTCTGGGT 993

RESULT 10

E28454

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

E28454
Hyaluronate synthase promoter DNA.

E28454

E28454.1 GI:13018346

JP 1999196875-A/2.

Mus sp.

E28454 2102 bp DNA linear PAT 18-JUN-2001

ORGANISM	Mus sp.	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2102)	
AUTHORS	Yoichi, Y., Naoki, I. and Koji, K.	
TITLE	Hyaluronate synthase promoter DNA	
JOURNAL	Patent: JP 1999196875-A 2 27-JUL-1999;	
COMMENT	SEIKAGAKU KOGYO CO LTD	
OS	Mus sp. (mouse)	
PN	JP 1999196875-A/2	
PD	27-JUL-1999	
PF	14-JAN-1998 JP 1998006191	
PR		
PI	YOICHI YAMADA, NAOKI ITANO, KOJI KIMATA	
PC	C12N15/09, C12N9/00, C12Q1/68// (C12N15/09, C12R1:91), C12N15/00, (C12N15/00, C12R1:91)	
CC	Strandedness: Double;	
CC	Topology: Linear;	
FH	Key	
FT	Location/Qualifiers	
FEATURES	source	
	1..2102	
	/organism="Mus sp."	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10095"	
ORIGIN		
	Query Match 63.3%; Score 678.2; DB 6; Length 2102;	
	Best Local Similarity 84.5%; Pred. No. 1e-80;	
	Matches 793; Conservative 0; Mismatches 128; Indels 18; Gaps 2;	
Qy	10 CAGACGGCCCAAGCCACTCTGACGCCGCGCTGCTCGGCGCTGGCCGAGGGTG 69	
Db		
Qy	55 CAGACATGCGCAAGCCCTCAGAGCAGCGGTTGCTGCTCGGCTGGCCGAGCGACA 114	
Db		
Qy	70 CTGACCATCGCTTCGCCCTCTATCTCGGCTCATGACCTGGGCTTACGCGCGCGG 129	
Db		
Qy	115 CTCAGATCATCTTTGGCCCTCATCTGGGCTCATGACCTGGGCTACGCGCAGGC 174	
Db		
Qy	130 GTGCGCTGGCTTCGATCGCTACGGCTCTCGGCTTCGCGCTTACGCGCGCTTCTT 189	
Db		
Qy	175 GTTCTCTGGCTTCAGATCGCTATGAGCTCTGCGCTTTGGCTTATGGGCTATCTC 234	
Db		
Qy	190 TCAGGGACCTGTGGCGCAGAGCTCTTCGGTACCTGGAGCAGCGCGGCTGGCGGC 249	
Db		
Qy	235 AGCGCACACCTAGTGGCAGAGCTCTTCTGCTTACCTGGAGCACCAGAGGTTGGCAG 294	
Db		
Qy	250 GCGGCGCG-----GGGCGCGCTGGATGCAGCCACCGCGCGCAGTGTGGCG 294	
Db		
Qy	295 GCTGGCGCGCGCTCTTTGGCGAAGGGGCGCTGGATGGGCGCACTGCACGCGGTGGCA 354	
Db		
Qy	295 CTGACCATCTCCGCTACACAGAGGAGCCCGCGCTACCTGGCCAGTGCCTGGCGCGCC 354	
Db		
Qy	355 CTCACCATCTCAGCTTACCAAGAGATCCCGCTTACCTGCGCCAGTGTTCACCTCGCG 414	
Db		
Qy	355 CGCGCGCTGTGTACCGCGCGCGGCTGGCGCTGCGCTCATGCTGGTGGATGGCAACGCG 414	
Db		
Qy	415 CGCGCGCTGTGTGTACCGCGCACACAGAGTTACGCGCTCATGCTGGTGGAGCGCAACGCG 474	
Db		
Qy	415 GCGGAGGACCTCTACATGCTGCATGTTCCGCGAGGTTCTTCGCTGCAGAGGACCCCGCG 474	
Db		
Qy	475 GCTGAGGATCTGTATATGTTGATCATGTTCCGAGAGTCTTCGCGGATGAGGACCCCGCG 534	
Db		
Qy	475 ACGTACGTGTGGGACGGCAATACCAACAGCCCTGGGAACCCCGCGCG-----CGGGCGCG 531	
Db		
Qy	535 ACTTATGTGTGGATGGCAACTACCATCAGCCCTGGGAACCCAGCGAGGCTACGCGGCT 594	
Db		
Qy	532 GTGGCGCGGAGCTATCGGAGAGGTGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAG 591	
Db		
Qy	595 GTCGGTGAAGTGTCTACCGGAGGTGAGCGGAGGACCCCGCGCGTGTGGCGGTGGAG 654	
Db		
Qy	592 GCGCTGTGTGAGACTCGCAGGTGCTGTGCTGCGCAGCGCTGGGGCGGCAAGCGCGAG 651	
Db		

Db	655 GCGCTGTGTGAGAACACACGCGGTGCGTGTGCGTGCAGCGCTTGGGGCGGCAACGTCGAG 714	
Qy	652 GTCATGTACACAGAGCTTCAAGGCGCTCGGAGATTTCGGTGGACTAGTCGACGCTCTGTGAC 711	
Db		
Db	715 GTCATGTACACAGAGCTTCAAGGCGACTCGGCGACTCCGTTGGACTACGTCAGGCTCTGTGAC 774	
Qy	712 TCGGACACAGAGTTGGACCCCATCGCACTGCTGAGAGCTCGTGGGCTACTCGACGAGGAC 771	
Db		
Db	775 TCAGACACAGACTAGACCCCATGCACTGCTGAGAGCTTGGAGTGTGGAGTGTGGATGAAGAC 834	
Qy	772 CCCCGGTAGGGGCTGTTGGTGGGAGCTGCGGATCCCTTAACCCCTCTGGACTCTCTGGGTC 831	
Db		
Db	835 CCCCGGTAGGGGCTGTTGGAGGGATGTGAGGATCCCTTAACCCCTCTGGACTCTCTGGGTC 894	
Qy	832 AGCTTCTTACGACGCTTGGGATCTGCTGAGTCTGCTTCAATGTGGAGCGGCTTGTTCAGAG 891	
Db		
Db	895 AGCTTCTTGGAGCAGTCTTCGATCTGCTGAGTACGCTTCAATGTGGAACGAGCTTGTTCAGAGC 954	
Qy	892 TACTTCCACTGTGTATCTGTCATCAGCGGTTCTCTAGGT 930	
Db		
Db	955 TACTTCCACTGTGTGTCTGTCATCAGTGTCTCTGGGT 993	
RESULT 11		
E30971		
LOCUS	E30971 2102 bp DNA linear PAT 18-JUN-2001	
DEFINITION	Hyaluronic acid synthetase modified protein.	
ACCESSION	E30971	
VERSION	E30971.1 GI:13017286	
KEYWORDS	JP 2000004886-A/1.	
SOURCE	Mus sp.	
ORGANISM	Mus sp.	
REFERENCE	1 (bases 1 to 2102)	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Naoki, I., Mamoru, Y. and Koji, K.	
TITLE	Hyaluronic acid synthetase modified protein	
JOURNAL	Patent: JP 2000004886-A 1 11-JAN-2000;	
COMMENT	SEIKAGAKU KOGYO CO LTD	
OS	Mus sp. (mouse)	
PN	JP 2000004886-A/1	
PD	11-JAN-2000	
PF	24-JUN-1998 JP 1998193788	
PR		
PI	NAOKI ITANO, MAMORU YOSHIDA, KOJI KIMATA	
PC	C12N15/09, C12N9/00, C12N15/00	
CC	Strandedness: Double;	
CC	Topology: Linear;	
FH	Key	
FT	Location/Qualifiers	
FEATURES	source	
	1..2102	
	/organism="Mus sp."	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10095"	
ORIGIN		
	Query Match 63.3%; Score 678.2; DB 6; Length 2102;	
	Best Local Similarity 84.5%; Pred. No. 1e-80;	
	Matches 793; Conservative 0; Mismatches 128; Indels 18; Gaps 2;	
Qy	10 CAGACGGCCCAAGCCACTCTGACGCCGCGCTGCTCGGCGCTGGCCGAGGGTG 69	
Db		
Qy	55 CAGACATGCGCAAGCCCTCAGAGCAGCGGTTGCTGCTCGGCTGGCCGAGCGACA 114	
Db		
Qy	70 CTGACCATCGCTTCGCCCTCTATCTCGGCTCATGACCTGGGCTTACGCGCGCGG 129	
Db		
Qy	115 CTCAGATCATCTTTGGCCCTCATCTGGGCTCATGACCTGGGCTACGCGCAGGC 174	
Db		
Qy	130 GTGCGCTGGCTTCGATCGCTACGGCTCTCGGCTTCGCGCTTACGCGCGCTTCTT 189	
Db		
Qy	175 GTTCTCTGGCTTCAGATCGCTATGAGCTCTGCGCTTTGGCTTATGGGCTATCTC 234	
Db		
Qy	190 TCAGGGACCTGTGGCGCAGAGCTCTTCGGTACCTGGAGCAGCGCGGCTGGCGGC 249	
Db		
Qy	235 AGCGCACACCTAGTGGCAGAGCTCTTCTGCTTACCTGGAGCACCAGAGGTTGGCAG 294	
Db		
Qy	250 GCGGCGCG-----GGGCGCGCTGGATGCAGCCACCGCGCGCAGTGTGGCG 294	
Db		
Qy	295 GCTGGCGCGCGCTCTTTGGCGAAGGGGCGCTGGATGGGCGCACTGCACGCGGTGGCA 354	
Db		
Qy	295 CTGACCATCTCCGCTACACAGAGGAGCCCGCGCTACCTGGCCAGTGCCTGGCGCGCC 354	
Db		
Qy	355 CTCACCATCTCAGCTTACCAAGAGATCCCGCTTACCTGCGCCAGTGTTCACCTCGCG 414	
Db		
Qy	355 CGCGCGCTGTGTACCGCGCGCGGCTGGCGCTGCGCTCATGCTGGTGGATGGCAACGCG 414	
Db		
Qy	415 CGCGCGCTGTGTGTACCGCGCACACAGAGTTACGCGCTCATGCTGGTGGAGCGCAACGCG 474	
Db		
Qy	415 GCGGAGGACCTCTACATGCTGCATGTTCCGCGAGGTTCTTCGCTGCAGAGGACCCCGCG 474	
Db		
Qy	475 GCTGAGGATCTGTATATGTTGATCATGTTCCGAGAGTCTTCGCGGATGAGGACCCCGCG 534	
Db		
Qy	475 ACGTACGTGTGGGACGGCAATACCAACAGCCCTGGGAACCCCGCGCG-----CGGGCGCG 531	
Db		
Qy	535 ACTTATGTGTGGATGGCAACTACCATCAGCCCTGGGAACCCAGCGAGGCTACGCGGCT 594	
Db		
Qy	532 GTGGCGCGGAGCTATCGGAGAGGTGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAG 591	
Db		
Qy	595 GTCGGTGAAGTGTCTACCGGAGGTGAGCGGAGGACCCCGCGCGTGTGGCGGTGGAG 654	
Db		
Qy	592 GCGCTGTGTGAGACTCGCAGGTGCTGTGCTGCGCAGCGCTGGGGCGGCAAGCGCGAG 651	
Db		

QY	190	TCAGCGCACCTGCTGGCGAGAGCCTCTTCGGCTACCTGGAGCACCGCGCGGTGGCGGG	249
Db	235	AGCGCACACCTAGTAGGCGACAGAGCCTCTTCGCTTACCTGGAGCACCGAGGTTGGCAGCG	294
QY	250	GGCGCGCG-----GGGGCGCTGGATGACAGCCACCGCGCGCAGTGTGGCG	294
Db	295	GCTGCGCGCGCTCTCTTGGCGAAGGGGCCCTTGGATGGCCCACTGACGCGAGCGTGGCA	354
QY	295	CTGACCATCTCCGCTTACAGAGAGACCCCGGTACTCTGCGCCAGTGTGGCTCCGCC	354
Db	355	CTCACCATCTCAGCCTTACCAAGAGGATCCCGCTTACCTGGCGCAGTGTGACCTCCGCG	414
QY	355	CGCGCCCTCTGTACCGCGCGCGGCTGCGGCTCTCATGGTGGTGGATGGCAACGCC	414
Db	415	CGCGCCTTGTCTGTACCGGCACAGAGGTTACGCGTCTCATGGTGGTGGCAACCGC	474
QY	415	GCGGAGGACCTTACATGGTGTGACATGTTCCGCGAGGTTCTTCGCTGACGAGACCCCGCC	534
Db	475	GCTGAGGATCTGTACATGGTGGACATGTTCCGAGAAAGTCTTCGCGCATGAGACCCCGCC	534
QY	475	ACGTACGTGTGGAGCGGCACTACCAACGCGCTGGGNAACCCGCGCGG---CGGGCGG	531
Db	535	ACTTATGTGTGGATGGCAACTACATCAGCCCTGGGAAACAGCGGAGGCTACCGGCGCT	594
QY	532	GTGCGCGCGGAGCCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGCGAGTGGAG	591
Db	595	GTGCGTGAAGTGTCTTACCGGAGGAGTGGAGCGGAGGAGACCCCGCGGCTTGGCGTGGAG	654
QY	592	CGCTGTGTGAGGACTCGCAGGTGTGTGTGCGCGCAGCGCTGGGGCGGCAAGCGCAG	651
Db	655	CGCTGTGTGAGAACACGCGAGGTGTGTGTGCGCTCAGCGCTTGGGGCGGCAACGTTGAG	714
QY	652	GTGATGTACAGCGCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGAC	711
Db	715	GTGATGTACAGAGCTTTCAGGSCATGGGCGACTCCGCTGAGTCTACGTGCAGGTCTGTGAC	774
QY	712	TCGACACAAAGTGTGGACCCCATGCTGCTGAGGCTCGTGGCGGTACTTGGACGAGGAC	771
Db	775	TCAGACACAAAGCTAGACCCCATGCTGCTGAGGCTGTGGAGCTTGTGGAGTGTGGATGAAGAC	834
QY	772	CCCCGGGTAGGGGCTGTGTGGTGGGAGACGTGCGGATTCCTTAAACCTCTCGGACTCTCTGGGTC	831

PR NAOKI ITANO, KOJI KIMATA
PI C12N15/09, C12N5/10, G01N33/50//C12N9/00, C12Q1/68, (C12N15/09, PC
PC C12R1:91),
PC (C12N5/10, C12R1:91), C12N15/00, C12N5/00, (C12N15/00, C12R1:91),
PC (C12N5/00, C12R1:91)
CC
FH Key Location/Qualifiers
FT CDS 49..1800.
1..2102
Location/Qualifiers
source
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

FEATURES
ORIGIN

Query Match 63.3%; Score 678.2; DB 6; Length 2102;
Best Local Similarity 84.5%; Pred. No. 1e-80; Mismatches 128; Indels 18; Gaps 2;
Matches 793; Conservative 0;

QY 10 CAGGACGCGCCCAAGCCCACTCTCTGACGCCCGCGCTGCTCGGCGCTGCGCGGAGGTTG 69
Db 55 CAGGACATGCCAAAGCCCTCAGAGGACGCGGTTGCTGCTCTGGCTTGGCCAGGCGAGCA 114
QY 70 CTGACCATCGCCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCCCGCGG 129
Db 115 CTCAGATCATCTTTGCCCTGCTCATCTGGGCTCATGACCTGGGCTTACGCCCGAGG 174
QY 130 GTGCGCTGGCTCCGATCGCTAGGCGCTCTGGGCTCTGGGCTCTAGCGGCGCTTCCTT 189
Db 175 GTTCCTCTGGCTTCAGATCGCTATGGACTCTCTGGGCTTGGGCTCTATGGGCGCATTCCTC 234
QY 190 TCAGCGCACCTGTGGCGCAGAGCCTCTTCGCTACCTTGGAGCACCGCGGCTTGGCGG 249
Db 235 AGCGCACACCTAGTAGGCGACAGGCTCTTCGCTTACCTTGGAGCACCGAAGGTTGGCAGCG 294
QY 250 GCGGCGG-----GGGGCGCTGGATGACGACCGCGCGGCAAGTGTGGG 294
Db 295 GCTGCGCGCGCTCTCTTGGCGAAGGGGCCCTTGGATGGCGGCACTGACGCGAGCGTGGCA 354
QY 295 CTGACCATCTCGGCTACCGAGGAGACCCCGCTACCTGCGGAGTGTGGTGGTGGCAACCGC 354
Db 355 CTCACCATCTCAGCCTTACCAAGAGGATCCCGCTTACCTTGGCGGAGTGTGACCTCCGCG 414
QY 355 CGCGCCTCTGTGTACCGCGCGCGCTGGCGGCTCTCATGCTGTGGTGGTGGTGGCAACCGC 414
Db 415 CGCGCTTGTGTGTACCGCGCACACGAGGTTACGCGTGTCTCATGCTGTGGTGGTGGCAACCGC 474
QY 415 GCGGAGGACCTCTACATGGTTCGACATGTTCCGCGAGGTTCTTCGCTGACGAGACCCCGCC 474
Db 475 GCTGAGGATCTGTACATGGTGGACATGTTCCGAGAAAGTCTTCGCGCATGAGACCCCGCC 534
QY 475 ACGTACGTGTGGAGCGGCACTACCAACGCGCTGGGNAACCCGCGCGG---CGGGCGG 531
Db 535 ACTTATGTGTGGATGGCAACTACATCAGCCCTGGGAAACAGCGGAGGCTACCGGCGCT 594
QY 532 GTGCGCGCGGAGCCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGCGAGTGGAG 591
Db 595 GTGCGTGAAGTGTCTTACCGGAGGAGTGGAGCGGAGGAGACCCCGCGGCTTGGCGTGGAG 654
QY 592 CGCTGTGTGAGGACTCGCAGGTGTGTGTGCGCGCAGCGCTGGGGCGGCAAGCGCAG 651
Db 655 CGCTGTGTGAGAACACGCGAGGTGTGTGTGCGGCTCAGCGCTTGGGGCGGCAACGTTGAG 714
QY 652 GTGATGTACAGCGCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGAC 711
Db 715 GTGATGTACAGAGCTTTCAGGSCATGGGCGACTCCGCTGAGTCTACGTGCAGGTCTGTGAC 774
QY 712 TCGACACAAAGTGTGGACCCCATGCTGCTGAGGCTCGTGGCGGTACTTGGACGAGGAC 771
Db 775 TCAGACACAAAGCTAGACCCCATGCTGCTGAGGCTGTGGAGCTTGTGGAGTGTGGATGAAGAC 834
QY 772 CCCCCGGGTAGGGGCTGTGTGGTGGGAGACGTGCGGATTCCTTAAACCTCTCGGACTCTCTGGGTC 831
Db 835 CCCCCGGGTAGGGGCTGTGTGGAGGGAGTGTGAGGATCTTAAACCTCTGAGACTCTCTGGGTC 894
QY 832 AGCTTCTTAAGCAGCGCTCGGATPACTGGGTAGCTTCAATGTGGAGCGGCTTGTGACAGG 891
Db 895 AGCTTCTTGAGCAGCTTTCGATCTAGGTAGGCTTCAATGTGGAGCGGCTTGTGACAGG 954
QY 892 TACTTCCATGTGTATCTGCAATCAGCGGTTCTCTAGGT 930
Db 955 TACTTCCATGTGTCTGTCATCAGTGTCTCTGGGT 993

RESULT 12
E34326
LOCUS 2102 bp DNA linear PAT 31-JAN-2002
DEFINITION DNA for gene targeting hyaluronic acid synthase gene.
ACCESSION E34326
VERSION E34326.1 GI:18624311
KEYWORDS JP 2000116382-A/1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2102)
Itano, N. and Kimata, K.
REFERENCE
AUTHORS DNA for gene targeting hyaluronic acid synthase gene
JOURNAL Patent: JP 2000116382-A 1 25-APR-2000;
SEIKAGAKU KOGYO CO LTD
OS Mus musculus (mouse)
PN JP 2000116382-A/1
PD 25-APR-2000
PF 13-OCT-1998 JP 1998291201

```

Db      835 CCCCCGGTAGGGCTGTTGGAGGGATGTGAGGATCCTTAACCTCTGGACTCTCGGTC 894
Qy      832 AGCTTCCTAAGCAGCGCTCGGACTACTGGGTAGCCTCAATGTGGAGCGGGCTTGTGAGAGC 891
Db      895 AGCTTCCTTGGACAGACTTTCGATACCTAGGAGTGGGTAGCCTCAATGTGGAGCGAGCTTGTGAGAGC 954
Qy      892 TACTTCCACTGTATCTCGCATCAGCGGTCTCTAGGT 930
Db      955 TACTTCCACTGTGTCTCGCATCAGTGTCTCTCGGGT 993

RESULT 13
MUSHAS
LOCUS      MUSHAS                2102 bp      mRNA      linear      ROD 06-FEB-1999
DEFINITION Mus musculus mRNA for hyaluronan synthase, complete cds.
ACCESSION D82964
VERSION    D82964.1 GI:1339939
KEYWORDS   hyaluronan synthase.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus

REFERENCE 1 (sites)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL    Itano,N. and Kimata,K.
MEDLINE    Expression cloning and molecular characterization of HAS protein, a
PUBMED     eukaryotic hyaluronan synthase
86215261   J. Biol. Chem. 271 (17), 9875-9878 (1996)
REFERENCE 2 (bases 1 to 2102)
AUTHORS    Itano,N.
JOURNAL    Unpublished
AUTHORS    Itano,N.
JOURNAL    Direct Submission
TITLE      Submitted (08-JAN-1996) Naoki Itano, Aichi Medical University,
JOURNAL    Institute for Molecular Science of Medicine, Nagakute, Aichi-gun,
Aichi 480-11, Japan (Tel:052-264-4811(ex.2087), Fax:0561-63-3532)
FEATURES   Location/Qualifiers
            1..2102
             /organism="Mus musculus"
             /mol_type="mRNA"
             /db_xref="taxon:10090"
             /cell_line="F3A H41"
             /clone_lib="cDNA/pcdNA1"
            49..1800
             /codon_start=1
             /product="hyaluronan synthase"
             /protein_id="BAAL1654.1"
             /db_xref="GI:1339940"
             /translation="MRQMPKPSBAARCCSLARRALTIIPALLILMTWYAAAGVP
LASDRYGLAFGLYGAFLSAHLVAQSLPAYLEHRRVAAARRSLAKGPLDAATARSVA
LTI5AYQDPAYLROCLTISARALLYPHTRLMVVDGNRAEDLYMDMFEVPAED
PATVYDNGHOPPEPATGAVGEVAREVEDPGRLEAVEALVTRRCVCVAORWG
KREVMYTFAPLSDVDYVQCDSDFLDPALLELVRLDPRVGNVGDVRLN
PLDSWVSLSLIRTFWAFNVERACQSFHCVSCISGPGLLYRNLLQOFLFANYNQK
LGTCTGDDRHLLNRMLSMGYATKYTSRSRCYSETSPFSFLRWLLWLLCVQGVALLAK
LYNALMWRHRAHWYEA VVSGLPFFVAATVRLRYAGRPWALLWLLCVQGVALLAK
AAFAWLRGLGSLVLYAPLYMCGLLIPAKFLALVTMNSQSGWGTSGRKKLAANYVP
LPLALWALLLGLGLARSVAQEARADWSGPSRAEAYHLAAGAGAYVAYVVMVLTIIYV
GVRLCRRRSRGYRVQV"

Query Match      63.3%; Score 678.2; DB 10; Length 2102;
Best Local Similarity 84.5%; Pred. No. 1e-80;
Matches 793; Conservative 0; Mismatches 128; Indels 18; Gaps 2;

Qy      10 CAGACGGCCCAAGCCACTCTCTGCAGCCCGCGCTCTCGGCTGGCCGAGGGTG 69
Db      55 CAGGACATGCCAAAGCCCTCAGAGCAGCGGCTGTGCTCTGCGCTGGCCGAGCGAGCA 114
Qy      70 CTGACCATCGCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTACGCCGCCGG 129
```

```

Db      115 CTACAGATCATCTTTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTACGCCGAGGC 174
Qy      130 GTGCGGCTGGGCTCCGATCGCTACGGGCTCTCTGGCCTTTCGGCCTTACCGGGGCTTCTT 189
Db      175 GTTCTCTGGCTTCAGATCGTATGAGATCTCTGGCCTTTCGGCCTTATATGGGGCATATCCTC 234
Qy      190 TCAGGGACCTGTGGGCGCAGAGCCTCTTGCCTACCTTGGAGACCCGCGGCTGGCGCG 249
Db      235 AGCGACACCTAGTGGCACAGAGCCTCTTCTGCTTACCTTGGAGCACCGAGGCTGGCAGCG 294
Qy      250 GCGGCGCG-----GGGCGCGCTGTGATGACGACCGCGCGCGCAGTGTGGCG 294
Db      295 GCTGCGCGCGCTCTTTGGGGAAGGGGCCCTGTGATGGGCCACTGACGCGAGGTGGCA 354
Qy      295 CTGACCATCTCCGCTTACAGAGAGACCCCGCTGACCTGCGCCAGTGTGCTGGCGTCCGCC 354
Db      355 CTACCATCTCAGCCTTACCAAGAGGATCCCGCTTACCTGCGCCAGTGTGCTTACCTCCGCG 414
Qy      355 CGCGCCTGTGTACCGCGCGCGCGCTGCGCGCTCTCATGTGTTGGATGGCAACCGC 414
Db      415 CGCGCCTGTGTACCGCGCACAGAGGTTACGCGTGTCTCATGTGTTGGAGCGGCAACCGC 474
Qy      415 GCCGAGGACCTCTACATGTGTGACATGTTCGCGAGGTCTTCGCTGACGAGGACCCCGCC 474
Db      475 GCTGAGGATCTGTATAGTGTGACATGTTCGAGAGAGTCTTCGCGGATGAGGACCCCGCC 534
Qy      475 ACGTACGTGTGGAGCGGCAATACACAGCCCTGGGAAACCCGCGCGG-----CGGCGCG 531
Db      535 ACTTATGTGGGATGGCACTACCATCAGCCCTGGGAAACCCGCGGAGGTACCGGGCGCT 594
Qy      532 GTGGCGCGGAGCCTATCGGAGGTGAGCGGAGGATCCTGCGCGCTGGCAGTGGAG 591
Db      595 GTCGCTGAAGTGCTTACCGGAGGTGAGCGGAGGACCCCGCGCGGCTTGGCGTGGAG 654
Qy      592 GCGCTGTGTGAGGACTCGCAGGTGCTGTGCGTGGCGAGCGCTGGGCGGCGCAAGCGAG 651
Db      655 GCGCTGTGTGAGAACACAGCAGGTGCTGTGCGTGGCTCAGCGTTCGGGCGGCAACGTGAG 714
Qy      652 GTCATGTACACAGCCTTCAAGCGCGCTCGAGAGATTCGGTGGACTACGTGACAGGTCTGTGAC 711
Db      715 GTCATGTACACAGCCTTTCAAGGCACTGGGCGGACTCCGTGGACTACGTGAGTCTGTGAC 774
Qy      712 TCGGACACAAAGTTGGACCCCATGGCAGCTGCTGAGAGCTCGTGGCGGCTACTGACAGAGAC 771
Db      775 TCAGACAAAGACTAGACCCCATGGCAGCTGCTGAGAGCTTGTCCAGAGTGTGGATGAAGAC 834
Qy      772 CCGCGGTAGGGGCTGTGTGGTGGGACGTGCGGATCCTTAACCCCTCTGGACTCTCGGTC 831
Db      835 CCGCGGTAGGGGCTGTGTGGAGGGGATGTAGGATCCTTAACCCCTCTGGACTCTCGGTC 894
Qy      832 AGCTTCCTAAGACCGCTCGCATCTGCGGTAGCCTTCAATGTGGAGCGGGCTTGTGACAGC 891
Db      895 AGCTTCCTTGGACAGCTCTTCGATCTAGCTAGCCTTCAATGTGGAAACGAGCTTGTGAGAGC 954
Qy      892 TACTTCCACTGTGTATCTCGCATCAGCGGTCTCTAGGT 930
Db      955 TACTTCCACTGTGTCTCGCATCAGTGTCTCTCGGT 993

RESULT 14
AC130783/c
LOCUS      AC130783                193986 bp      DNA      linear      HTG 19-NOV-2002
DEFINITION Pan troglodytes clone CH251-426A12, WORKING DRAFT SEQUENCE, 8
ordered pieces.
ACCESSION AC130783
VERSION    AC130783.2 GI:25100968
KEYWORDS   HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
REFERENCE   1 (bases 1 to 193986)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
```

AUTHORS

Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Carliaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Masiello,C., Maskeri,B., McDaniel,J.,
Padurigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.

TITLE

NISC Comparative Sequencing Initiative

REFERENCE

2 (bases 1 to 193986)

AUTHORS

Green,E.D.

TITLE

Direct Submission

JOURNAL

Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717

JOURNAL

Grovmont Circle, Gaithersburg, MD 20877, USA

REFERENCE

3 (bases 1 to 193986)

AUTHORS

Green,E.D.

TITLE

Direct Submission

JOURNAL

Submitted (19-NOV-2002) NIH Intramural Sequencing Center, 8717

JOURNAL

Grovmont Circle, Gaithersburg, MD 20877, USA

COMMENT

On Nov 19, 2002 this sequence version replaced gi:22218453.

COMMENT

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.zoo@nhgri.nih.gov

----- Project Information

Center project name: dxd

Center clone name: 426A12

The sequence data in this record represents an 'enhanced' version of a phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 191764 bases at least Q40

Consensus quality: 192510 bases at least Q30

Consensus quality: 192998 bases at least Q20

Insert size: 184000; agarose-fp

Insert size: 193286; sum-of-contigs

Quality coverage: 11.49x in Q20 bases; agarose-fp

Quality coverage: 10.94x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 11512: contig of 11512 bp in length

* 11513 11612: gap of unknown length

* 11613 59488: contig of 47876 bp in length

* 59489 59588: gap of unknown length

* 59589 103266: contig of 43678 bp in length

* 103267 103366: gap of unknown length

* 103367 127885: contig of 24519 bp in length

* 127886 127985: gap of unknown length

* 127986

* 129692: contig of 1707 bp in length

* 129693

* 129792: gap of unknown length

* 129793

* 131077: contig of 1285 bp in length

* 131078

* 131177: gap of unknown length

* 131178

* 174663: contig of 43486 bp in length

* 174664

* 193986: gap of unknown length.

* 174764

* 193986: contig of 19223 bp in length.

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

* 193986

```

Db      132408  CGCGCCGAGGACCTCTACATGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCC 132349
QY      472     GCCAGTACGTGTGGGACGGCAACTACACACAGCCCTCGGAAACCCGCGCGCGCGCGCGCG 531
Db      132348  GCCAGTACGTGTGGGACGGCAACTACACACAGCCCTCGGAAACCCGCGCGCGCGCGCGCGCG 132289
QY      532     GTGGGCGCGGAGCCATCTCGGAGGTGAGCGGAGGATCTCTGGGCGCTGCGCAGTGGAG 591
Db      132288  GTGGGCGCGGAGCCCTACCGGAGGTGGAGCGGAGGATCTCTGGGCGCTGCGCAGTGGAG 132229
QY      592     GCGCTGTGTGAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCAG 651
Db      132228  GCGCTGTGTGAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCAG 132169
QY      652     GTCATGTACACAGCCCTTCAAGCGCTCGGAGATTGCGTGGACTGACGTGCAGGTCTGT 708
Db      132168  GTCATGTACACAGCCCTTCAAGCGCTCGGAGATTGCGTGGACTGACGTGCAGGTGAGT 132112

RESULT 15
AC137058/c
LOCUS
DEFINITION AC137058 185623 bp DNA linear HTG 21-FEB-2003
Papio anubis clone RP41-126M5, WORKING DRAFT SEQUENCE, 10 ordered
pieces.
ACCESSION AC137058
VERSION AC137058.3 GI:28460766
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Papio anubis (olive baboon)
ORGANISM Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Papio.
1. (bases 1 to 185623)
Akter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Hachichi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Latic,P., Lee-Lin,S.-Q., Legaapi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Mastello,C., Maskeri,B., McDownell,J.,
Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Weherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2. (bases 1 to 185623)
Green,E.D.
Direct Submission
Submitted (15-NOV-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
3. (bases 1 to 185623)
Green,E.D.
Direct Submission
Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
On Feb 21, 2003 this sequence version replaced gi:27476124.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoonhgri.nih.gov
----- Project Information
Center project name: dtj
Center clone name: 126M05

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average

coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 183092 bases at least Q40

Consensus quality: 183851 bases at least Q30

Consensus quality: 184434 bases at least Q20

Insert size: 160000; agarose-fp

Insert size: 184723; sum-of-contigs

Quality coverage: 14.13x in Q20 bases; agarose-fp

Quality coverage: 12.24x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 10 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 9984: contig of 9984 bp in length

* 9985 10084: gap of unknown length

* 10085 49805: contig of 39721 bp in length

* 49806 49905: gap of unknown length

* 49906 66936: contig of 17031 bp in length

* 66937 67036: gap of unknown length

* 67037 83604: contig of 16568 bp in length

* 83605 83704: gap of unknown length

* 83705 108246: contig of 24542 bp in length

* 108247 108346: gap of unknown length

* 108347 112021: contig of 3675 bp in length

* 112022 112121: gap of unknown length

* 112122 134817: contig of 22696 bp in length

* 134818 134917: gap of unknown length

* 134918 147290: contig of 12373 bp in length

* 147291 147390: gap of unknown length

* 147391 185337: contig of 37947 bp in length

* 185338 185437: gap of unknown length

* 185438 185623: contig of 186 bp in length.

FEATURES

```

source
1. .185623
   /organism="Papio anubis"
   /mol_type="genomic DNA"
   /db_xref="taxon:9555"
   /clone="RP41-126M5"
   /clone_lib="RP41"
1. .15263
   /note="clone overlaps with GenBank Accession Number
AC130273 clone RP41-30706 (center project name dau)"
misc_feature
1. .9984
   /note="assembly_fragment
vector_side:left"
misc_feature
10085. .49805
   /note="assembly_fragment"
misc_feature
49906. .66936
   /note="assembly_fragment"
misc_feature
67037. .83604
   /note="assembly_fragment"
misc_feature
83705. .108246
   /note="assembly_fragment"
misc_feature
108347. .112021
   /note="assembly_fragment"
misc_feature
112122. .134817
   /note="assembly_fragment"
misc_feature
134918. .147290
   /note="assembly_fragment"
misc_feature
147391. .185337

```


/note="assembly_fragment"
185438..185623
/note="assembly_fragment
clone_end:SP6
vector_side:right"

misc_feature

ORIGIN

```
Query Match      59.3%; Score 635.4; DB 2; Length 185623;
Beat Local Similarity 94.3%; Pred. No. 1.4e-75;
Matches 676; Conservative 0; Mismatches 26; Indels 15; Gaps 1;

Qy 7 CAGCAGGAGCGCCCAAGCCACTCTGCAGCCGCGCTGCTCGGCTGCGCCGCGAGG 66
Db 138595 CAGCGGAGACACCCCAAGCCACTCTGCAGCCGCGCTGCTCGGCTGCGCCGCGAGG 138536

Qy 67 GTGCTGACCATCGCCTTGGCCCTGTCTCATCTCTGGCCCTCATGACCTGGCCCTTACGCCGCC 126
Db 138535 GTGCTGACCATCGCCTTGGCCCTGTCTCATCTCTGGCCCTCATGACCTGGCCCTTACGCCGCC 138476

Qy 127 GGGGTGCGCGTGGCTTCCGATCGCTACCGGCTCTCTGGCTTCTGGCTCTACCGGGCCTTC 186
Db 138475 GGGGTGCGCGTGGCTTCCGATCGCTACCGGCTCTCTGGCTTCTGGCTCTACCGGGCCTTC 138416

Qy 187 CTTTCAGCGCACCTGGTGGCGGAGAGCCTCTTCGGCTGACCTGGAGCACCGGCGGGTGGCG 246
Db 138415 CTCTCGGCGCACCTGGTGGCGGAGAGCCTCTTCGGCTGACCTGGAGCACCGGCGGGTGGCG 138356

Qy 247 GCGGCGCGG-----CGGGGCGCGCTGGATGACGACCGCGCGCAGTGTG 291
Db 138355 GCGGCGCGGCGGCGCGCGGCGCACGGGGCGCTTGGATGACGACCGCGCGCAGCGTG 138296

Qy 292 GCGCTGACCATCTCGCCTTACAGAGGACCCCGCTACCTGCGCCAGTGCCTGCGCTCC 351
Db 138295 GCGCTGACCATCTCGCCTTACAGAGGACCCCGCTACCTGCGCCAGTGCCTGCGCTCC 138236

Qy 352 GCCCGGCGCCTGCTGTACCGCGCGCGCGCTGCGCGCTCTCATGGTGGTGGATGGCAAC 411
Db 138235 GCCCGGCGCCTGCTGTACCGCGCGCGCGCTGCGCGCTCTCATGGTGGTGGATGGCAAC 138176

Qy 412 CGCGCGGAGGACCTCTACATGGTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCC 471
Db 138175 CGCGCGGAGGACCTCTACATGGTGACATGTTCCGCGAGGTCTTCGCGAGGACCCC 138116

Qy 472 GCCACGTACGTGTGGGACGGCAACTTACACAGCCCTGGGAACCCGCGCGCGCGCGCG 531
Db 138115 GCCACGTACGTGTGGGACGGCAACTTACACAGCCCTGGGAACCCGCGCGCGCGCGCG 138056

Qy 532 GTGGCGCGGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAG 591
Db 138055 GTGGCGCGGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAG 137996

Qy 592 GCGCTGGTGAGGACTCGCAGGTGCTGTGCGGCGCAGCGCTGGGGCGGCAAGCGCGAG 651
Db 137995 GCGCTGGTGAGGACTCGCAGGTGCTGTGCGGCGCAGCGCTGGGGCGGCAAGCGCGAG 137936

Qy 652 GTCATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAGGCTGTGT 708
Db 137935 GTCATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAGGCTGTGT 137879
```

Search completed: March 13, 2005, 08:01:29
Job time : 5065.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 02:13:47 ; Search time 5115.08 Seconds
(without alignments)
10259.279 Million cell updates/sec

Title: US-10-672-399-5

Perfect score: 1083

Sequence: 1 atgagacagcaggacgcgc.....tccaagtgcctactccgtga 1083

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.on.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	933	86.1	2088	9 HSU59269	U59269 Human hyalu
2	933	86.1	2116	6 AR220003	AR220003 Sequence
3	920.8	85.0	2119	9 BC035837	BC035837 Homo sapi
4	920.6	85.0	2108	9 HUMHAS	D84424 Homo sapien
5	920.6	85.0	2117	6 AR137038	AR137038 Sequence
6	855	78.9	2109	9 AY463695	AY463695 Papio anu
7	697.2	64.4	229155	9 AC018755	AC018755 Homo sapi
8	679.4	62.7	2095	10 AB097568	AB097568 Rattus no
9	677.8	62.6	2102	6 E13681	E13681 DNA encodin
10	677.8	62.6	2102	6 E28454	E28454 Hyaluronate
11	677.8	62.6	2102	6 E30971	E30971 Hyaluronon
12	677.8	62.6	2102	6 E34326	E34326 DNA for gen
13	677.8	62.6	2102	10 MUSHAS	D82964 Mus musculu
14	669	61.8	193986	2 AC130783	AC130783 Pan trogl
15	635.4	58.7	185623	2 AC137058	AC137058 Papio anu
16	620.4	57.3	1581	6 CQ731079	CQ731079 Sequence
17	492	45.4	293184	2 AC079487	AC079487 Mus muscu
18	490.8	45.3	72955	2 AC108651	Continuation (4 of
19	490.8	45.3	257325	2 AC116203	AC116203 Rattus no

BV099313 RPAMMSQ0
 BV160501 RPAMMSQ0
 AY437407 Danio rer
 AB055979 Oryctolag
 BV161377 RPAMMSQ0
 CQ723590 Sequence
 AF234839 Homo sapi
 AF232772 Homo sapi
 CQ776471 Sequence
 AB159675 Sub scrof
 BC047963 Xenopus l
 M22249 X.laervis DG
 AB097569 Rattus no
 AF015780 Xenopus l
 BC077983 Xenopus l
 E30973 Hyaluronon
 AR373269 Sequence
 CQ777380 Sequence
 UB6408 Mus musculu
 CQ776472 Sequence
 BC021853 Homo sapi
 BV161384 RPAMMSQ0
 BX539313 Zebrafish
 BV160507 RPAMMSQ0
 BV161386 RPAMMSQ0
 BV099322 RPAMMSQ0

ALIGNMENTS

RESULT 1
 HSU59269
 LOCUS Human hyaluronan synthase mRNA, complete cds.
 DEFINITION
 ACCESSION U59269
 VERSION U59269.1 GI:1556464
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 2088)
 AUTHORS Shyjan,A.M., Heidin,P., Butcher,E.C., Yoshino,T. and Briskin,M.J.
 TITLE Functional cloning of the cDNA for a human hyaluronan synthase
 JOURNAL J. Biol. Chem. 271 (38), 23395-23399 (1996)
 MEDLINE
 PUBMED 8798544
 REFERENCE 2 (bases 1 to 2088)
 AUTHORS Briskin,M.J. and Shyjan,A.M.
 TITLE Direct Submission
 JOURNAL Submitted (24-MAY-1996) LeukoSite Inc., 215 First Street,
 Cambridge, MA 02142, USA

FEATURES

source
 1..2088
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 36..1772
 /codon_start=1
 /product="hyaluronan synthase"
 /protein_id="AAC50706.1"
 /db_xref="GI:1556465"
 /translation="MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAGV
 PLASDRYGLLAFLGFLGAPLSAHLVAQSLFAYLEHRRVAAAARGPLDAAATASVALTIS
 AYQDPRYLRLQCLASARALYPRARLVLYVDGNRAEDLYMVFREVFADDEPATY
 VWDGNQHPWEPAAGAGVAGEAEDEPGLRVALRVCVCAVQWGGGKREV
 MYTFKALGDSVDYVQVCDSDTRDPMALLELVLDVLDPRVAGVGGVRLINPLDSW
 VFLSLRYMVAFNVERACQSYFHCVSCISGCLGLYRNLLQQLFEAWYKQFLGTHC
 TFGDRLHNTNRLMSNGYATKTSRSRCVSETPSSFLRWLSQQTNRKSKYFRELWYNAL
 WWHRHAWMTYEAVVSGLFPFPFVALVRLFYAGRPWALLWVLLCVQCAAKAAFAA
 WLRCRLRLLSLYAPLYMCLGFLPAKFLALVTMNSQGWGTSGRKLAANYPLPLAL
 WLLLLLGLVRSVAHEARADWSGSPRAAEAYHLAAGAYVGVYVWVAMLTLYVWVGVRRL

ORIGIN		CRRRTGGYRVQV"	
Query Match		86.1%; Score 933; DB 9; Length 2088;	
Best Local Similarity		90.6%; Pred. No. 3.6e-114;	
Matches 1048; Conservative		0; Mismatches 35; Indels 74; Gaps 2;	
Qy	1	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCCCGCCGCTGCTCGGGCTGGCC	60
Db	36	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCCCGCCGCTGCTCGGGCTGGCC	95
Qy	61	CGGAGGGTGCTGACCATCGCCTTCGGCCCTGCTCATCTTGGGCTCATGACCTGGGCTTAC	120
Db	96	CGGAGGGTGCTGACCATCGCCTTCGGCCCTGCTCATCTTGGGCTCATGACCTGGGCTTAC	155
Qy	121	GCCGCCGGGGTGCCGCTGCGCTCCGATGCTACGCGCTCTTGGGCTTGGGCTTACGCGG	180
Db	156	GCCGCCGGGGTGCCGCTGCGCTCCGATGCTACGCGCTCTTGGGCTTGGGCTTACGCGG	215
Qy	181	GCCTTCTTTTCAGCGCACCTGGTGCGCAGAGCTCTTTCGCGTACCTTGGAGCACCGCGCG	240
Db	216	GCCTTCTTTTCAGCGCACCTGGTGCGCAGAGCTCTTTCGCGTACCTTGGAGCACCGCGCG	275
Qy	241	GTGGCGCGCGCGCGCGCGCGCGCGCTGGATGCGACCAACCGCGCGCAGTGTGGCGCTGACC	300
Db	276	GTGGCGCGCGCGCGCGCGCGCGCGCTGGATGCGACCAACCGCGCGCAGTGTGGCGCTGACC	335
Qy	301	ATCTTCGCGCTTACAGAGGAGACCCCGCTACCTGCGCAGTGCCTTGGCGCTCGCCCGCGCC	360
Db	336	ATCTTCGCGCTTACAGAGGAGACCCCGCTACCTGCGCAGTGCCTTGGCGCTCGCCCGCGCC	395
Qy	361	CTGCTGTACCCGCGCGCGCGCTGCGGCTCCTCATGGTGTGTGATGCGAACCGCGCGCGAG	420
Db	396	CTGCTGTACCCGCGCGCGCGCTGCGGCTCCTCATGGTGTGTGATGCGAACCGCGCGCGAG	455
Qy	421	GACCTCTACATGTGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGCTAC	480
Db	456	GACCTCTACATGTGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGCTAC	515
Qy	481	GTGTGGAGCGGCAACTACCAACGAGCCCTGGGAAACCCCGCGCGCGCGCGGTGGGGGCC	540
ORIGIN		Query Match 86.1%; Score 933; DB 6; Length 2116;	
		Best Local Similarity 90.6%; Pred. No. 3.6e-114;	
		Matches 1048; Conservative 0; Mismatches 35; Indels 74; Gaps 2;	
Qy	1	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCCCGCCGCTGCTCGGGCTGGCC	60
Db	36	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCCCGCCGCTGCTCGGGCTGGCC	95
Qy	61	CGGAGGGTGCTGACCATCGCCTTCGGCCCTGCTCATCTTGGGCTCATGACCTGGGCTTAC	120
Db	96	CGGAGGGTGCTGACCATCGCCTTCGGCCCTGCTCATCTTGGGCTCATGACCTGGGCTTAC	155
Qy	121	GCCGCCGGGGTGCCGCTGCGCTCCGATGCTACGCGCTCTTGGGCTTTCGGGCTTACGCGG	180
Db	156	GCCGCCGGGGTGCCGCTGCGCTCCGATGCTACGCGCTCTTGGGCTTTCGGGCTTACGCGG	215
Qy	181	GCCTTCTTTTCAGCGCACCTGGTGCGCAGAGCTCTTTCGCGTACCTTGGAGCACCGCGCG	240
Db	216	GCCTTCTTTTCAGCGCACCTGGTGCGCAGAGCTCTTTCGCGTACCTTGGAGCACCGCGCG	275
Qy	241	GTGGCGCGCGCGCGCGCGCGCGCTGGATGCGACCAACCGCGCGCAGTGTGGCGCTGACC	300
Db	276	GTGGCGCGCGCGCGCGCGCGCGCTGGATGCGACCAACCGCGCGCAGTGTGGCGCTGACC	335
Qy	301	ATCTTCGCGCTTACAGAGGAGACCCCGCTACCTGCGCAGTGCCTTGGCGCTCGCCCGCGCC	360
Db	336	ATCTTCGCGCTTACAGAGGAGACCCCGCTACCTGCGCAGTGCCTTGGCGCTCGCCCGCGCC	395
Qy	361	CTGCTGTACCCGCGCGCGCGCTGCGGCTCCTCATGGTGTGTGATGCGAACCGCGCGCGAG	420
Db	396	CTGCTGTACCCGCGCGCGCGCTGCGGCTCCTCATGGTGTGTGATGCGAACCGCGCGCGAG	455
Qy	421	GACCTCTACATGTGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGCTAC	480
Db	456	GACCTCTACATGTGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGCTAC	515
Qy	481	GTGTGGAGCGGCAACTACCAACGAGCCCTGGGAAACCCCGCGCGCGCGGTGGGGGCC	540
ORIGIN		Query Match 86.1%; Score 933; DB 6; Length 2116;	
		Best Local Similarity 90.6%; Pred. No. 3.6e-114;	
		Matches 1048; Conservative 0; Mismatches 35; Indels 74; Gaps 2;	
Qy	1	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCCCGCCGCTGCTCGGGCTGGCC	60
Db	36	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCCCGCCGCTGCTCGGGCTGGCC	95
Qy	61	CGGAGGGTGCTGACCATCGCCTTCGGCCCTGCTCATCTTGGGCTCATGACCTGGGCTTAC	120
Db	96	CGGAGGGTGCTGACCATCGCCTTCGGCCCTGCTCATCTTGGGCTCATGACCTGGGCTTAC	155
Qy	121	GCCGCCGGGGTGCCGCTGCGCTCCGATGCTACGCGCTCTTGGGCTTTCGGGCTTACGCGG	180
Db	156	GCCGCCGGGGTGCCGCTGCGCTCCGATGCTACGCGCTCTTGGGCTTTCGGGCTTACGCGG	215
Qy	181	GCCTTCTTTTCAGCGCACCTGGTGCGCAGAGCTCTTTCGCGTACCTTGGAGCACCGCGCG	240
Db	216	GCCTTCTTTTCAGCGCACCTGGTGCGCAGAGCTCTTTCGCGTACCTTGGAGCACCGCGCG	275
Qy	241	GTGGCGCGCGCGCGCGCGCGCGCTGGATGCGACCAACCGCGCGCAGTGTGGCGCTGACC	300
Db	276	GTGGCGCGCGCGCGCGCGCGCGCTGGATGCGACCAACCGCGCGCAGTGTGGCGCTGACC	335
Qy	301	ATCTTCGCGCTTACAGAGGAGACCCCGCTACCTGCGCAGTGCCTTGGCGCTCGCCCGCGCC	360
Db	336	ATCTTCGCGCTTACAGAGGAGACCCCGCTACCTGCGCAGTGCCTTGGCGCTCGCCCGCGCC	395
Qy	361	CTGCTGTACCCGCGCGCGCGCTGCGGCTCCTCATGGTGTGTGATGCGAACCGCGCGCGAG	420
Db	396	CTGCTGTACCCGCGCGCGCGCTGCGGCTCCTCATGGTGTGTGATGCGAACCGCGCGCGAG	455
Qy	421	GACCTCTACATGTGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGCTAC	480
Db	456	GACCTCTACATGTGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGCTAC	515
Qy	481	GTGTGGAGCGGCAACTACCAACGAGCCCTGGGAAACCCCGCGCGCGCGGTGGGGGCC	540
Db	516	GTGTGGAGCGGCAACTACCAACGAGCCCTGGGAAACCCCGCGCGCGCGGTGGGGGCC	575
Qy	541	GGAGCCTATCGGAGGTGAGGCGGAGGATCTTGGCGGCTTGGCAGTGTGGAGGCGCTGGTG	600
Db	576	GGAGCCTATCGGAGGTGAGGCGGAGGATCTTGGCGGCTTGGCAGTGTGGAGGCGCTGGTG	635
Qy	601	AGGACTCGCAGGTGCTGTGCGTGGCGCAGCGCTTGGCGCGGCAAGCGCGAGTTCATGTAC	660
Db	636	AGGACTCGCAGGTGCTGTGCGTGGCGCAGCGCTTGGCGCGGCAAGCGCGAGTTCATGTAC	695
Qy	661	ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGGTCTGTGACTCGGACACA	720
Db	696	ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGGTCTGTGACTCGGACACA	755
Qy	721	AGGTGACCCCATGGCACTGCTGGAGCTGCTGGGGTACTGGACGAGGACCCCGCGGTA	780
Db	756	AGGTGACCCCATGGCACTGCTGGAGCTGCTGGGGTACTGGACGAGGACCCCGCGGTA	815
Qy	781	GGGCTGTGTGGGGAGCTGCGGATCTTAACCTCTGGAATCTTGGGCTGAGTTCCTTA	840
Db	816	GGGCTGTGTGGGGAGCTGCGGATCTTAACCTCTGGAATCTTGGGCTGAGTTCCTTA	875
Qy	841	AGCAGCTGCGGATCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTTCAGAGCTACTTCCAC	900
Db	876	AGCAGCTGCGGATCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTTCAGAGCTACTTCCAC	935
Qy	901	TGTGTATCTCTGATCAGGGTCTCTAG-----	928
Db	936	TGTGTATCTCTGATCAGGGTCTCTAG-----	995
Qy	929	-----AATCTGCGCAGGCGCCCGAGGAGCAGCGGATGATGCC	965

ALLLGLVRSVAHEARDWSGSPRAAEAVHLAAGAGAYVGYVMAMLTLYWVGVRRLC
RRRTGYRVQV"

ORIGIN

Query Match 85.0%; Score 920.8; DB 9; Length 2119;
Best Local Similarity 90.3%; Pred. No. 1.5e-112;
Matches 1037; Conservative 0; Mismatches 37; Indels 74; Gaps 2;

Qy 10 CAGACGGCCCAAGCCACTCTCGACCGCCGCGCTGCTCGGCGCTGCGCGGAGGGTG 69
Db 48 CAGACGGCCCAAGCCACTCTCGACCGCCGCGCTGCTCGGCGCTGCGCGGAGGGTG 107

Qy 70 CTGACCATCGCTTTCGCGCTCTCATCTCGGCGCTCATGACCTGCGCGCTAGCGCGCGGG 129
Db 108 CTGACCATCGCTTTCGCGCTCTCATCTCGGCGCTCATGACCTGCGCGCTAGCGCGCGGG 167

Qy 130 GTGCGCGTGGCTCCGATCGCTACGCGCTCTCGGCTTTCGCGCTTACGCGGCGCTTCTT 189
Db 168 GTGCGCGTGGCTCCGATCGCTACGCGCTCTCGGCTTTCGCGCTTACGCGGCGCTTCTT 227

Qy 190 TCAGGCACTGGTGGCGACAGCTCTTCGCTACCTGGAGCACCGCGGCTGCGCGG 249
Db 228 TCAGGCACTGGTGGCGACAGCTCTTCGCTACCTGGAGCACCGCGGCTGCGCGG 287

Qy 250 GCGCGCGGGCGCGCTGGATGACGACCGCGCGCAGTGTGGCGCTGACCATCTCGGCC 309
Db 288 GCGCGCGGGCGCGCTGGATGACGACCGCGCGCAGTGTGGCGCTGACCATCTCGGCC 347

Qy 310 TACGAGGAGACCCCGGTACTCTCGCGAGTGTCTGGCGTCCGCGCGCGCTCTGTAC 369
Db 348 TACGAGGAGACCCCGGTACTCTCGCGAGTGTCTGGCGTCCGCGCGCGCTCTGTAC 407

Qy 370 CCGCGCGCGCGCTGCGGCTCTCATGTTGTGTGATGCGCAACCGCGCGGAGCTCTAC 429
Db 408 CCGCGCGCGCGCTGCGGCTCTCATGTTGTGTGATGCGCAACCGCGCGGAGCTCTAC 467

Qy 430 ATGTTGACATGTTCCGCGAGTCTTCGCTACGAGGACCCCGCACGCTAGTGTGGAC 489
Db 468 ATGTTGACATGTTCCGCGAGTCTTCGCTACGAGGACCCCGCACGCTAGTGTGGAC 527

Qy 490 GGCAACTACCAACGAGCTTGGGAACCCCGCGCGCGCGCGGTGGCGCGGAGCCAT 549
Db 528 GGCAACTACCAACGAGCTTGGGAACCCCGCGCGCGCGGTGGCGCGGAGCCAT 587

Qy 550 CCGAGGTGGAGCGGAGGATCTCTGGGCGGTGCGAGTGGAGCGCTGGTGGAGCTCGC 609
Db 588 CCGAGGTGGAGCGGAGGATCTCTGGGCGGTGCGAGTGGAGCGCTGGTGGAGCTCGT 647

Qy 610 AGTGGCTGTGCTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTGATGACACAGCCTTC 669
Db 648 AGTGGCTGTGCTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTGATGACACAGCCTTC 707

Qy 670 AAGCGCTCGGAGATTCGGTGGACTACGTCAGGTCTGTGATCTCGGACACAAAGTTGGAC 729
Db 708 AAGCGCTCGGAGATTCGGTGGACTACGTCAGGTCTGTGATCTCGGACACAAAGTTGGAC 767

Qy 730 CCATGCGACTGCTGGAGCTCTGGGCTGCTGAGCGAGGACCCCGGCTAGGGCTGTT 789
Db 768 CCATGCGACTGCTGGAGCTCTGGGCTGCTGAGCGAGGACCCCGGCTAGGGCTGTT 827

Qy 790 GGTGGGACGTGCGGATCTTAACTCTGAGCTCTGGGCTGAGCTTCTTAAAGCAGCTG 849
Db 828 GGTGGGACGTGCGGATCTTAACTCTGAGCTCTGGGCTGAGCTTCTTAAAGCAGCTG 887

Qy 850 CGATCTCGGCTAGCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCACTGTGTATCC 909
Db 888 CGATCTCGGCTAGCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCACTGTGTATCC 947

Qy 910 TGCATGAGCGTCTCTAG----- 928
Db 948 TGCATGAGCGTCTCTAGGAGCTATAGGAATAACCTTTGACGAGTCTTTTGGAGCC 1007
Qy 929 -----AATCCTGCCAGGCCCCCAGGAGCAGCGGATGATGCCCTCATTCCT 974

Db 1008 TGGTACACCAAGAGTTCTCTGGGTACCCACTGTACTTTTGGGATGACCGGCACCTCACC 1067
Qy 975 CGCCCC-----CGTGAGGTACACCTCCAGGTCCCGTCTGTACTCA 1015
Db 1068 AACCGATGTCTCAGCATGGGTATGCTACCAAGTACACCTCCAGGTCCCGTCTGTACTCA 1127
Qy 1016 GAGACGCCCTCGCTCTCTCGGCTGCTGAGCGAGCAGACAGCTGCTCCAGTCTGTAC 1075
Db 1128 GAGACGCCCTCGCTCTCTCGGCTGCTGAGCGAGCAGACAGCTGCTCCAGTCTGTAC 1187
Qy 1076 TTCCGTCA 1083
Db 1188 TTCCGTCA 1195

RESULT 4

HUMAS LOCUS 2108 bp mRNA linear PRI 06-FEB-1999
DEFINITION Homo sapiens mRNA for hyaluronan synthase, complete cds.
ACCESSION D84424
VERSION D84424.1 GI:1401033
KEYWORDS hHAS; hyaluronan synthase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Itano, N. and Kimata, K.
Molecular cloning of human hyaluronan synthase
JOURNAL Biochem. Biophys. Res. Commun. 222 (3), 816-820 (1996)
MEDLINE 96244584
PUBMED 8651928
REFERENCE 2 (bases 1 to 2108)
AUTHORS Itano, N.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2108)
AUTHORS Itano, N.
TITLE Direct Submission
Submitted (16-APR-1996) Naoki Itano, Aichi Medical University,
Institute for Molecular Science of Medicine; Nagakute, Aichi-gun,
Aichi 480-11, Japan (Tel:052-264-4811(ex.2087), Fax:0561-63-3532)

FEATURES

source

1..2108
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="brain"
/clone_lib="1gt11"
/dev_stage="fetal"
1..2108
/genes="hHAS"
149..1780
/genes="hHAS"
/codon_start=1
/product="hyaluronan synthase"
/protein_id="BA012351.1"
/db_xref="GI:1401034"

gene

CDS

ORIGIN

Query Match 85.0%; Score 920.6; DB 9; Length 2108;
Best Local Similarity 90.2%; Pred. No. 1.5e-112;
Matches 1038; Conservative 0; Mismatches 39; Indels 74; Gaps 2;

QY 706 TGTGACTCGGACACAGGTTGGACCCCATGGACCTGCTGGAGCTCGTGGGGTACTGGAC 765
Db 756 TGTGACTCGGACACAGGTTGGACCCCATGGACCTGCTGGAGCTCGTGGGGTACTGGAT 815
QY 766 GAGACACCCCGGTAGGGGCTGTTGGTGGGACGTCGGATCCTTAACCCCTCGACTCC 825
Db 816 GAGACACCCCGGTAGGGGCTGTTGGTGGGACGTCGGATCCTTAACCCCTCGACTCC 875
QY 826 TGGGTGAGCTTCTTAAGCAGCGCTCGGATACCTGAGTGGGTAGCTTCAATGTGGAGCGGCTTGT 885
Db 876 TGGGTGAGCTTCTTAAGCAGCGCTCGGATACCTGAGTGGGTAGCTTCAATGTGGAGCGGCTTGT 935
QY 886 CAGAGCTACTTCCACTGTGTATCTCGCATCAGCGTCTCTAG 928
Db 936 CAGAGCTACTTCCACTGTGTCTCTCGCATCAGTGTCTCTAG 978

RESULT 7
AC018755/c
LOCUS AC018755 229155 bp DNA linear PRI 26-JUL-2000
DEFINITION Homo sapiens chromosome 19, BAC BC330783 (CIT-HSPC_470E3), complete
sequence.
AC018755
VERSION AC018755.3 GI:9454515
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.
1 (bases 1 to 229155)
Lamerdin, J.E., McCready, P.M., Skowronski, E., Viswanathan, V.,
Burkhardt-Schultz, K., Gordon, L., Dias, J., Ramirez, M., Stilwagen, S.,
Phan, H., Velasco, N., Do, L., Regalla, W., Terry, A., Brower, A.,
Garnes, J., Danganan, L., Erler, A., Christensen, M., Georgescu, A.,
Avila, J., Liu, S., Andrise, T., Frankheim, M., Attix, C.,
Amico-Keller, G., Coefield, J., Duarte, S., Lucas, S., Bruce, R.,
Thomas, P., Quan, G., Kronmiller, B., Arellano, A., Sanders, C., Ow, D.,
Nolan, M., Trong, S., Kobayashi, A., Olsen, A.S. and Carrano, A.V.
Sequence analysis of a 5-Mb region in 19q13.4
Unpublished
2 (bases 1 to 229155)
DOE Joint Genome Institute.
Direct Submission
Submitted (19-DEC-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 229155)
Lamerdin, J.E.
Direct Submission
Submitted (26-JUL-2000) DOE Joint Genome Institute, Lawrence
Livermore National Laboratory, Livermore, CA 94550
On Jul 26, 2000 this sequence version replaced gi:7458780.
Map and sequence oriented from centromere to q-telomere. BC330783
(CIT-HSPC_470E3) is overlapped on the left by BC849408
(CITB-El_3073N11, AC020914) from bases 1 to 125 of this accession,
and overlaps cosmid R28782 (LNL-R_248F10, AC005946) on the right
from bases 217,905 to 229,155 of this accession. Additional
chromosome 19 map and sequence information may be obtained at:
<http://www-bio.llnl.gov/bbrp/genome/genome.html>.

FEATURES
Location/Qualifiers
1..229155
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
/clone="CITC-470E3 (LNL clone name BC330783)"
/clone_lib="Caltech human BAC library C"
201..262
/rpt_family="L2"
<299..16870
/gene="ZNF175"
note="zinc finger protein 175; synonym: OTK18"
join(<299..550,8540..8666,8936..9031,13776..15616)

CDS
/gene="ZNF175"
/product="ZNF175 mRNA"
join(479..550,8540..8666,8936..9031,13776..15616)
/gene="ZNF175"
/codon_start=1
/product="zinc finger protein 175"
/protein_id="AAF87843.1"
/db_xref="GI:9454517"
translation="MPADVNLISQKQPVQLGPEQDSCBASVFEVDVDFSGREWQOL
DPAQCLYRDVMLELYSHLPAVGVIHPNPEVIFMLKEKEPRVEAEVSHQRCORREF
GLEIPQKEISKKASFOQDMVGEFTRDGSWCSILELRDLADRTKKDEQNOQPMHSHA
FFNKKTINTSNCBYKDPGKMIRTPHLLASQKQPKCCLFESLKLNLVNGQNESN
DTEQLDDVWGSGQLFSHSSDACSKNTHGETFCGKQCRVKQSLKQHQIHTOK
KIDGSCGSGSFTQKSHLFAQQRHSHVGNLHCEGCKGKAFMPQLKSLYLDHTDIP
CICEKCVFIORSELLTHQTRKPKYKCHDCGKAFOMLSLFRHORTHREKLYE
CSECGKGFSONSLIHKQHTGEROVACSECGKAFQKSTLSLHQRTHSGOKSYVCI
ECQAFIQKXHLIWHQSHSGEKEPYQCHNCCKSFISQSLDIHRIHTGERPYECSDC
GKTFTQSHLNIHQIHTGERHHVCSGCKAFNOKSILSHQRIHTGERPKYKCEGCK
AFTSKQFQTHQTHMGEKPYVTECGKAFNGSNFHKQIHTHTRPPFVYKCGKAF
VQSELTHTHRTMGEKPYECLDCGKSFSSKKPQLKQVHQRHTHTRPPVYKCGKAFN
RSNFKHTHTHTRDYSKCSYKVGFTKQ"
1097..1198
/rpt_family="MER99"
1559..1829
/rpt_family="AluJb"
complement(2130..2355)
/rpt_family="L1ME"
complement(2359..2627)
/rpt_family="AluY"
complement(2816..3183)
/rpt_family="HAL1"
complement(3290..3598)
/rpt_family="AluSg"
complement(3611..3760)
/rpt_family="MIR"
complement(3766..4067)
/rpt_family="HAL1"
4098..4215
/rpt_family="AluJb"
complement(4339..4406)
/rpt_family="HAL1"
complement(4518..4724)
/rpt_family="HAL1"
complement(4803..5110)
/rpt_family="AluY"
5266..5561
/rpt_family="AluSg"
5635..5932
/rpt_family="AluJo"
5984..6120
/rpt_family="AluJb"
6121..6423
/rpt_family="AluSx"
6424..6579
/rpt_family="AluJb"
6647..6913
/rpt_family="L1ME"
6936..7295
/rpt_family="MER33"
7700..7785
/rpt_family="L2"
complement(9216..9434)
/rpt_family="L1ME"
complement(9435..9745)
/rpt_family="AluY"
complement(9746..9815)
/rpt_family="L1ME"
complement(10009..10303)
/rpt_family="AluSp"
11783..12637
/rpt_family="L1MB2"
12180..12330
/gene="ZNF175"
STS

/standard_name="SHGC-32377"
13436..13619
/rpt_family="MIR"
16068..16112
/rpt_family="(TA)n"
complement(16113..16329)
/rpt_family="MER46A"
16330..16390
/rpt_family="(TA)n"
16403..16439
/rpt_family="(TATG)n"
16793..17013
/rpt_family="MERS8A"
17024..17052
/rpt_family="(CA)n"
17311..17545
/rpt_family="MIR"
complement(18683..18896)
/rpt_family="MIR"
complement(18915..19183)
/rpt_family="AluY"
19336..19599
/rpt_family="MIR"
19811..19928
/rpt_family="GA-rich"
22181..23081
/rpt_family="L2"
complement(23692..23901)
/rpt_family="L2"
complement(23902..24201)
/rpt_family="AluSg1"
complement(24202..24651)
/rpt_family="L2"
26136..26434
/rpt_family="LTR12"
26435..26525
/rpt_family="LTR30"
26526..26643
/rpt_family="LTR12"
complement(26686..26884)
/rpt_family="AluSg/x"
27065..27240
/rpt_family="MIR"
complement(27238..27554)
/rpt_family="L2"
28406..28433
/rpt_family="GC-rich"
28641..28678
/rpt_family="GC-rich"
complement(28997..29259)
/rpt_family="MSR1"
complement(30219..30489)
/rpt_family="L2"
30753..30882
/rpt_family="FLAM_C"
31138..31416
/rpt_family="AluY"
complement(31547..31716)
/rpt_family="MER3"
32041..32112
/rpt_family="L1MEC"
32117..32516
/rpt_family="MLT2E"
35056..35118
/rpt_family="MIR"
complement(35578..35636)
/rpt_family="MERSA"
complement(35690..35820)
/rpt_family="MIR"
35872..35902
/rpt_family="(TG)n"
complement(35906..36209)
/rpt_family="AluJo"

Query Match	64.4%	Score 697.2;	DB 9;	Length 229155;
Best Local Similarity	99.6%	Pred. No. 1.2e-83;		
Matches 699;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

Qy	7	CAGCAGGACGCGCCCAAGCCCACTCTCTGACAGCCCGCGCTGCTCCGGCTGCGCCGAGG 66
Db	147050	CAGCAGGACGCGCCCAAGCCCACTCTCTGACAGCCCGCGCTGCTCCGGCTGCGCCGAGG 146991
Qy	67	GTGCTGACATGCGCTTTCGCTGCTCATCTGCGGCTCATAGACTGGGCTACGCGCC 126
Db	146990	GTGCTGACATGCGCTTTCGCTGCTCATCTGCGGCTCATAGACTGGGCTACGCGCC 146931
Qy	127	GGGTCGCGCTGCGCTCCGATCGCTAGGCGCTCTGCGCTTCGGCTCTACGGGCGCTTC 186
Db	146930	GGGTCGCGCTGCGCTCCGATCGCTAGGCGCTCTGCGCTTCGGCTCTACGGGCGCTTC 146871
Qy	187	CTTTACGCGCACCTGTGGCGCAGAGCTCTCTCGCTACCTGGAGCACCGCGCGGTGGCG 246
Db	146870	CTTTACGCGCACCTGTGGCGCAGAGCTCTCTCGCTACCTGGAGCACCGCGCGGTGGCG 146811
Qy	247	GCGGCGCGCGGGCGCTTGGATGAGCACCGCGCGCAGTGTGGCGTGACCATCTCC 306
Db	146810	GCGGCGCGCGGGCGCTTGGATGAGCACCGCGCGCAGTGTGGCGTGACCATCTCC 146751
Qy	307	GCCTACGAGGAGNACCCGCGGTACTGCGCAGTGTGGCGCTGCGCTCCGCGCGCTGCTG 366
Db	146750	GCCTACGAGGAGNACCCGCGGTACTGCGCAGTGTGGCGCTGCGCTCCGCGCGCTGCTG 146691
Qy	367	TACCGCGCGCGGGCTGCGCTCTCATGTGTGATGGCAACCGCGCGCGAGACCTC 426
Db	146690	TACCGCGCGCGGGCTGCGCTCTCATGTGTGATGGCAACCGCGCGCGAGACCTC 146631
Qy	427	TACATGTCGACATGTTCCCGAGGTCTTCGCTGACGAGACCCCGCACGTACGTGTGG 486
Db	146630	TACATGTCGACATGTTCCCGAGGTCTTCGCTGACGAGACCCCGCACGTACGTGTGG 146571
Qy	487	GACGGCAACTACCACGAGCCCTGGGACCCGCGCGCGCGCGCGGTGGCGCGAGCC 546
Db	146570	GACGGCAACTACCACGAGCCCTGGGACCCGCGCGCGCGCGGTGGCGCGAGCC 146511
Qy	547	TATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGGCGCTGTGAGACT 606
Db	146510	TATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGGCGCTGTGAGACT 146451
Qy	607	CGCAGGTGCTGTGCGTGCGCGCAGCGCTGGGGCGCGCAAGCGCGAGGTCTATGACAGCC 666
Db	146450	CGCAGGTGCTGTGCGTGCGCGCAGCGCTGGGGCGCGCAAGCGCGAGGTCTATGACAGCC 146391
Qy	667	TTCAAGCGCTCGGAGATTTCGGTGGACTACGTGCAGTCTGT 708
Db	146390	TTCAAGCGCTCGGAGATTTCGGTGGACTACGTGCAGTCTGT 146349

RESULT 8	AB097568	2095 bp	mrna	linear	ROD 28-APR-2004
LOCUS	Rattus norvegicus	HAS1	RNA for hyaluronan synthase 1, complete cds.		
DEFINITION	AB097568	GI:26453348			
VERSION	AB097568.1				
KEYWORDS	Rattus norvegicus (Norway rat)				
SOURCE	Rattus norvegicus				
ORGANISM	Rattus norvegicus				
REFERENCE	1				
AUTHORS	Itano, N., Sawai, T., Atsumi, F., Miyaishi, O., Taniguchi, S., Kannagi, R., Hamaguchi, M. and Kinata, K.				
TITLE	Selective expression and functional characteristics of three Mammalian hyaluronan synthases in oncogenic malignant transformation				

J. Biol. Chem. 279 (18), 18679-18687 (2004)
14724275
REFERENCE
2 (bases 1 to 2095)
Itano,N.
Direct Submission
Submitted (05-DEC-2002) Naoki Itano, Aichi Medical University, Aichi
Institute for Molecular Science of Medicine, Nagakute, Aichi, Aichi
480-1195, Japan (E-mail: itano@amugw.aichi-med-u.ac.jp,
Tel:81-52-264-4811(ex.2095), Fax:81-561-63-3532)
Location/Qualifiers
1..2095
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1..2095
/gene="HAS1"
43..1794
/gene="HAS1"
/codon_start=1
/product="hyaluronan synthase 1"
/protein_id="BAC43730.1"
/db_xref="GI:26453349"
/translation="NRQMPKPESEARCCSGIARRVLITITFALLILGLMTWAYAGVP
LASDYGGLALGLYGLSALHVAQSLFAYLEHRRVTVAAARFAPKAGDRAATARSVA
LTISAYQDPTYLQCLTSARALLYPRFLVLMVVDGNRAEDLVMDFREVFAD
PATYVMDGNYPQWPEAAGAVEGAYREVEADPGRLAVALVTRRCVCAQWVG
GKREVMYAFKALGSDVDVQVCDSDRLDPALDELVRDLDDPRVAGVGDVRLN
PLDSWVFLSLRYVWVAENVARACQSYFHCVSCISGPIGLVRLNLLQOFLEAWYNOKE
LGTCTCFDHRHLNRLMISGYATKYTSRSCYSETSPSFLRLNSQOTRWSKYPREK
LYNALWHRHAAWMTYEAIVSGLFPFFVAATVRLFYAGRPWALLWLLVCQGVALAK
AAFAAWRGCLLRVLLSLYPLMCGLLPAKFLALVLMVQSGWTSGRKLAANYVP
LPLALLMALLGLLIRLSVAQEVRAVDSGFSRAAEAYHLAGASAYVAYVWIMLTIVYV
GVRLCRRRSRGYRVQV"

Query Match 62.7%; Score 679.4; DB 10; Length 2095;
Best Local Similarity 84.6%; Pred. No. 1.1e-80;
Matches 793; Conservative 0; Mismatches 126; Indels 18; Gaps 2;
QY 10 CAGGACGGCCCAAGCCACTCTCCAGCCGCGCTCTCGGCTCGCCGCGAGGTG 69
DB 49 CAGGACATGCCAAGCCCTCAGAGCAGCAGCTTCTCGTCTCGGCTCTCGCCAGCGGTG 108
QY 70 CTGACCATCGCTTCGCCCTCTCATCTCGGCTCTCATGACCTGGGCTCTACGCCCGCGG 129
DB 109 CTCAGATCACCTTCGCCCTCTCATCTCGGCTCTCATGACCTGGGCTCTACGCCAGGA 168
QY 130 GTGCCGCTCGCTCGATCGCTACGCCCTCTCGGCTCTCGGCTCTACGCCGCTTCCTT 189
DB 169 GTACCTCTGGCTTCTGATCCCTATGCGCTCTCGGCTCTGGGCTCTATGGGCGTTCCTC 228
QY 190 TCAGGACACCTGGTGGCGCAGACCTCTTCGGTACCTGGAGCAGCCGCGGCGGTGGCGG 249
DB 229 AGTGCGCACCTAGTGGCAGAGCCTCTTCGGTACCTGGAGCAGCCGAGGAGGTGACCGGTG 288
QY 250 GCGGCGCG-----GCGGCGCTGGATGACGACCCAGCCGCGGAGTGTGGCG 294
DB 289 GCTGGCGCGCGCTTTTTCGAGAGGACCCCTGGATGGGACCATCGCGGAGCGTGGCA 348
QY 295 CTGACCATCTCGCTACAGAGAGACCCCGGCTACCTCGCGCAGTGTGGCGTGGCGTCCGCC 354
DB 349 CTCACCATCTCAGCTTACAGAGAGACCCCACTTACCTGGCGCAGTGTGACCTCCGCG 408
QY 355 CGCGCCCTCTGATACCGCGCGCGCTGGCGCTCTCATGGTGGTGGATGGCAACGC 414
DB 409 CGCGCTCTGCTGTACCGCGCAGCGGCTGGCGCTGTATGGTGTGAGCGGCAATCGC 468
QY 415 GCCGAGGACCTCTACATGCTGCACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCC 474
DB 469 GCGGAGGATCTGATCGTGGACATGTTCCGAGAGTCTTCGCGAGATGAGGACCTCGCC 528
QY 475 ACGTACGTGTGGGACGGCACTTACCAACAGCCCTCGGAAACCCGC---GCGCGCGGCGCG 531

Db 529 ACTTATGTGTGGATGGCAACTACCATCAGCCTTGGGAGCCAGCTGAGCGCGCGGTGCT 588
QY 532 GTGGCGCGCGAGACCTATCGGAGGTGAGGCGGAGGATCCTGGCGGCTGGCAGTGGAG 591
Db 599 GTGGGTGAAGTGCCTACCGCGAGGTGAGGCTGAGGACCTTGGCGGCTGGCGGTAGAG 648
QY 592 GCGCTGTGTGAGGACTCGCAGGTGCTGTGCGTGGCGCAGCCTGGGCGCGCAAGCGCGAG 651
Db 649 GCGCTGTGTGAGGACCCGCGAGGTGCTGTGCGTGGCTGACGCGTGGGCTGGCAAGCGCGAA 708
QY 652 GTCATGTATACACAGCCTTCAAGCGCTCGGAGNTTCGTTGGACTGCTGAGCTCTGTGAC 711
Db 709 GTCATGTATACAGCCTTCAAGGCACTCGGTGACTCCGTGGACTCATGTGAGGTCTGTGAC 768
QY 712 TCGGACACAAGGTTGGACCCCATGCACTGTCTGGAGCTCGTGGCGGTACTTGCACGAGGAC 771
Db 769 TCAGACACAAGGTTAGACCCCATGCACTGTCTGGAGCTTGTGCGAGTGTCTGATGAGAC 828
QY 772 CCGCGGTGTGGGCTGTGGTGGGACGTGGGATCCTTAAACCTCTTGAACCTCTTGGGTC 831
Db 829 CCGCGGTGTGGTGTGTTGGAGGACATGTGAGATCCTTAAACCTCTTGGACTCTTGGGTC 888
QY 832 AGCTTCTTAAAGCAGCTCGGATGCTGCTGAGCTTCAATGTGGAGCGGCTTGTGAGGC 891
Db 889 AGCTTCTTGAAGCAGCTTTCGATGCTGCTGAGCTTCAATGTGGAGCGGCTTGTGAGGC 948
QY 892 TACTTCCACTGTGATCTCGCATCAGCGCTCTCTAG 928
Db 949 TACTTCCACTGTGATCTCGCATCAGCGCTCTCTAG 985

RESULT 9
E13681
LOCUS
E13681
DEFINITION
DNA encoding novel mouse hyaluronate synthetase.
E13681
ACCESSION
E13681.1 GI:3252450
VERSION
JP 1997224674-A/1.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 2102)
AUTHORS
Itano,N. and Kimata,H.
TITLE
POLYPEPTIDE OF NEW HYALURONIC ACID-SYNTHETIC ENZYME AND DNA CODING
THE SAME
JOURNAL
Patent: JP 1997224674-A 1 02-SEP-1997;
COMMENT
OS Mus musculus (mouse)
PN JP 1997224674-A/1
PD 02-SEP-1997
PF 26-FEB-1996 JP 1996038336
PI ITANO NAKKI, KIMATA HIROHARU
PC C12N15/09,C07H21/04,C12N9/00//A61K48/00,C12N1/21,PC
(C12N9/00,
PC C12R1:91),(C12N9/00,C12R1:19),(C12N1/21,C12R1:19); CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..2102
FT /organism="Mus musculus"
FT /cell_line="FM3A"
FT 5'UTR 1..48
FT CDS 49..1800
FT /product="Novel mouse hyaluronate synthetase"
FT 3'UTR 1801..2102
FT Location/Qualifiers
1..2102
/organism="Mus musculus"
/mol_type="genomic DNA"

ORIGIN		/db_xref="taxon:10090"	
Query Match		62.6%;	Score 677.8; DB 6; Length 2102;
Best Local Similarity		84.5%;	Pred. No. 1.8e-80;
Matches		792; Conservative	0; Mismatches 127; Indels 18; Gaps 2;
Qy	10	CAGGACGGCCCAAGCCACTCTCTGACGCGCCGCTGCTCGGCGCTGCGCGGAGGGTG	69
Db	55	CAGGACATGCCAAAGCCCTCAGAGCAGCGGTTGCTCTTGGCGCTTGGCGGAGCA	114
Qy	70	CTGACCATCGCTTCCGCTCTGCTATCTCTGGGCTCATGACTGGGCTTACGCGCCGG	129
Db	115	CTCAGATCATCTTGGCTCTGCTATCTGGGCTCATGACTGGGCTTACGCGGAGC	174
Qy	130	GTGCGCTGGCTCGATCGCTACGCGCTCTGCGCTTGGCGCTTACGCGGCTTCTT	189
Db	175	GTTCCTCTGGCTTCAGATCGCTATGGACTCTGCGCTTGGCGCTTATGGGCTTCTC	234
Qy	190	TCAGCGCACCTTGGTGGCGCAGAGCTCTTGGCTTACCTGGAGCACCGCGGCTGGCGG	249
Db	235	AGCGCACACCTAGTGGCAGAGCTCTTGGCTTACCTGGAGCACCGAAGGCTGGCAGC	294
Qy	250	GCSSCGG-----GGGCGCGTGGATGAGCCAGCCAGCGCGCAGTGTGGCG	294
Db	295	GCTGCGGGCGCTCTTGGCGAAGGGGCCCTTGGATGGGCCCTGACGCGAGGTGGCA	354
Qy	295	CTGACCATCTCCGCTTACAGGAGGACCCCGCGTACCTGGCGCAGTGCTGGCGTCCGCC	354
Db	355	CTCACCATCTCAGCCTACCAAGAGGATCCCGCTTACCTGGCGCAGTGTGACCTCGCG	414
Qy	355	CGGCGCTTGTATACCGCGCGCGCTGGCGTCTCATGTGTGGTGGATGGCAACCGC	414
Db	415	CGGCGCTTGTATACCGCGCACGAGGTTACGCGTGTCTATGTGGTGGAGCGGCAACCGC	474
Qy	415	GCGGAGGACCTTACATGGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCC	474
Db	475	GCTGAGGATCTGTATACATGGTGGATGTTCCGAGAGTCTTCGCGGATGAGGACCCCGCC	534
Qy	475	ACGTACGTGGGACGGCAACTACCAAGCCCTGGGAAACCGCGCGG-----CGGCGCG	531
Db	535	ACTTATGTGGGATGGCAACTACCATCAGCCCTGGGAAACCGCGGCGCTACGCGGCT	594
Qy	532	GTGGGCGCGGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAG	591
Db	595	GTGCGTGAAGGTGCTACCGGAGGTGGAGCGGAGGACCCCGCGGCTTGGCGGTTGAG	654
Qy	592	GCGCTGTGAGGACTCGCAGGTGGGTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGGAG	651
Db	655	GCGCTGTGAGAACACGCGAGGTGGGTGTGCTGGCTCAGCGTTGGGGCGGCAAGCTGAG	714
Qy	652	GTCATGTACACAGCGCTTCAAGCGCTCGGAGATTGGTGGACTACGTGCAAGTCTGTGAC	711
Db	715	GTCATGTACACAGCTTTCAAGGCACTGGGCGACTCCGTTGGACTACGTGCAAGTCTGTGAC	774
Qy	712	TCGAGACAGAGTTGAGCCCATGGCACTGCTGGAGCTCGTGGGTTACTGGACGAGGAC	771
Db	775	TCAGACACAGACTAGACCCCATGGCACTGCTGGAGCTTGTGGAGTGTGGATGGAAGAC	834
Qy	772	CCCCGGGTAGGGCTGTGGTGGGACGCTGGGATCCCTTAACCTCTGGACTCTCGGGTC	831
Db	835	CCCCGGGTAGGGCTGTGGAGGGGATGTGAGGATCTTTAACCCTCTGGACTCTCGGGTC	894
Qy	832	AGCTTCTTACGACGCTGCGATACGTGGTACGCTTCAATGTGGACGCGGCTTGTGAGC	891
Db	895	AGCTTCTTACGAGCTTTCGATACGTGGTACGCTTCAATGTGGAAACGAGCTTGTGAGC	954
Qy	892	TACTTCCACTGTGTATCTCGCATCAGCGGTCTCTAG	928
Db	955	TACTTCCACTGTGTCTCGCATCAGTGGTCTCTCTGG	991
RESULT 10			

E28454	E28454	2102 bp	DNA	linear	PAT 18-JUN-2001
LOCUS	Hyaluronate synthase promoter DNA.				
DEFINITION	E28454				
ACCESSION	E28454.1	GI:13018346			
VERSION	JP 1999196875-A/2.				
KEYWORDS	Mus sp.				
SOURCE	Mus sp.				
ORGANISM	Mus sp.				
REFERENCE	1 (bases 1 to 2102)				
AUTHORS	Yoichi, Y., Naoki, I. and Koji, K.				
TITLE	Hyaluronate synthase promoter DNA				
JOURNAL	Patent: JP 1999196875-A 2 27-JUL-1999;				
COMMENT	SEIKAGAKU KOGYO CO LTD				
OS	Mus sp. (mouse)				
PN	JP 1999196875-A/2				
PD	27-JUL-1999				
PF	14-JAN-1998 JP 1998006191				
PR	YOICHI YAMADA, NAOKI ITANO, KOJI KIMATA				
PI	C12N15/09, C12N9/00, C12Q1/68// (C12N15/09, C12R1:91), C12N15/00,				
PC	(C12N15/00, C12R1:91)				
CC	Strandedness: Double;				
CC	Topology: Linear;				
EH	Key	Location/Qualifiers			
FT	CDS	49..1800.			
FEATURES		Location/Qualifiers			
source		1..2102			
ORIGIN		/organism="Mus sp."			
		/mol_type="genomic DNA"			
		/db_xref="taxon:10095"			
Query Match		62.6%;	Score 677.8; DB 6; Length 2102;		
Best Local Similarity		84.5%;	Pred. No. 1.8e-80;		
Matches		792; Conservative	0; Mismatches 127; Indels 18; Gaps 2;		
Qy	10	CAGGACGGCCCAAGCCACTCTCTGACGCGCCGCTGCTCGGCGCTGCGCGGAGGGTG	69		
Db	55	CAGGACATGCCAAAGCCCTCAGAGCAGCGGTTGCTCTTGGCGCTTGGCGGAGCA	114		
Qy	70	CTGACCATCGCTTCCGCTCTGCTCATCTGGGCTCATGACTGGGCTTACGCGCGGG	129		
Db	115	CTCAGATCATCTTGGCTCTGCTCATCTGGGCTCATGACTGGGCTTACGCGGAGC	174		
Qy	130	GTGCGCTGGCTCGATCGCTACGCGCTCTGCGCTTGGCGCTTACGCGGCTTCTT	189		
Db	175	GTTCCTCTGGCTTCAGATCGCTATGGACTCTGCGCTTGGCGCTTATGGGCTTCTC	234		
Qy	190	TCAGCGCACCTTGGTGGCGCAGAGCTCTTGGCTTACCTGGAGCACCGCGGCTGGCGG	249		
Db	235	AGCGCACACCTAGTGGCAGAGCTCTTGGCTTACCTGGAGCACCGAAGGCTGGCAGC	294		
Qy	250	GCSSCGG-----GGGCGCGTGGATGAGCCAGCCAGCGCGCAGTGTGGCG	294		
Db	295	GCTGCGGGCGCTCTTGGCGAAGGGGCCCTTGGATGGGCCCTGACGCGAGGTGGCA	354		
Qy	295	CTGACCATCTCCGCTTACAGGAGGACCCCGCGTACCTGGCGCAGTGCTGGCGTCCGCC	354		
Db	355	CTCACCATCTCAGCCTACCAAGAGGATCCCGCTTACCTGGCGCAGTGTGACCTCGCG	414		
Qy	355	CGGCGCTTGTATACCGCGCGCGCTGGCGTCTCATGTGTGGTGGATGGCAACCGC	414		
Db	415	CGGCGCTTGTATACCGCGCACGAGGTTACGCGTGTCTATGTGGTGGAGCGGCAACCGC	474		
Qy	415	GCGGAGGACCTTACATGGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCC	474		
Db	475	GCTGAGGATCTGTATACATGGTGGATGTTCCGAGAGTCTTCGCGGATGAGGACCCCGCC	534		
Qy	475	ACGTACGTGGGACGGCAACTACCAAGCCCTGGGAAACCGCGCGG-----CGGCGCG	531		
Db	535	ACTTATGTGGGATGGCAACTACCATCAGCCCTGGGAAACCGCGGCGCTACGCGGCT	594		
Qy	532	GTGGGCGCGGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAG	591		
Db	595	GTGCGTGAAGGTGCTACCGGAGGTGGAGCGGAGGACCCCGCGGCTTGGCGGTTGAG	654		
Qy	592	GCGCTGTGAGGACTCGCAGGTGGGTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGGAG	651		
Db	655	GCGCTGTGAGAACACGCGAGGTGGGTGTGCTGGCTCAGCGTTGGGGCGGCAAGCTGAG	714		
Qy	652	GTCATGTACACAGCGCTTCAAGCGCTCGGAGATTGGTGGACTACGTGCAAGTCTGTGAC	711		
Db	715	GTCATGTACACAGCTTTCAAGGCACTGGGCGACTCCGTTGGACTACGTGCAAGTCTGTGAC	774		
Qy	712	TCGAGACAGAGTTGAGCCCATGGCACTGCTGGAGCTCGTGGGTTACTGGACGAGGAC	771		
Db	775	TCAGACACAGACTAGACCCCATGGCACTGCTGGAGCTTGTGGAGTGTGGATGGAAGAC	834		
Qy	772	CCCCGGGTAGGGCTGTGGTGGGACGCTGGGATCCCTTAACCTCTGGACTCTCGGGTC	831		
Db	835	CCCCGGGTAGGGCTGTGGAGGGGATGTGAGGATCTTTAACCCTCTGGACTCTCGGGTC	894		
Qy	832	AGCTTCTTACGACGCTGCGATACGTGGTACGCTTCAATGTGGACGCGGCTTGTGAGC	891		
Db	895	AGCTTCTTACGAGCTTTCGATACGTGGTACGCTTCAATGTGGAAACGAGCTTGTGAGC	954		
Qy	892	TACTTCCACTGTGTATCTCGCATCAGCGGTCTCTAG	928		
Db	955	TACTTCCACTGTGTCTCGCATCAGTGGTCTCTCTGG	991		

Qy	532	GTGGGCGCCGAGCCTATCGGAGATGGAAGCGGAGGATCCTGGGCGGCTGGCAGTGGAG	591
Db	595	GTCCGTGAAGGTGCCTACCGGAGGTGGAGGCGGAGGACCCCGGCGGTTGGCGGTGGAG	654
Qy	592	CGCGTGGTGAAGCATCGCAGAGTGCCTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCGAG	651
Db	655	CGCGTGGTGAAGCAACGAGAGTGCCTGTGCGTGGCTCAGCGTTGGGGCGGCAACCGTGA	714
Qy	652	GTCATGTACACAGCCCTTCAAGGCGCTCGGAGATTCCGTGGGACTAGCTGCAGGCTCTGTGAC	711
Db	715	GTCATGTACACAGCTTTCAGGCACTGGGCGACTCCGTGGACTACGTGCAGGCTCTGTGAC	774
Qy	712	TCGGACACAAGGTTGGACCCCATGCGCATCTGTGAGAGTCTGTCGGGAGTCTGGACAGGAC	771
Db	775	TCAGACACAAGACTAGACCCCATGCGCACTGCTGTGAGCTTGTGCGAGTGTGTGATGAAGAC	834
Qy	772	CCCCGGGTAGGGCTGTGGTGGGGACGTGCGGATCCTTTAACCCCTCTGAGCTCCTGGGTC	831
Db	835	CCCCGGGTAGGGCTGTGGAGGGAGTGTGAGGATCCTTTAACCCCTCTGAGCTCCTGGGTC	894
Qy	832	AGCTTCTTAAGCAGCCTCGCATCTGCGTAGTGGTAGCCTTCAATGTGGAGCGGCTGTGACAGC	891
Db	895	AGCTTCTTGAAGCTTTCGATCTAGTGGTAGCCTTCAATGTGGAACGAGCTTGTGACAGC	954
Qy	892	TACTTCCACTGTATCCTGCATCAGCGGTCTCTAG	928
Db	955	TACTTCCACTGTGTCTCTGCATCAGTGGTCTCTCG	991
RESULT 11			
E30971	E30971	2102 bp	DNA
LOCUS	E30971	2102 bp	linear
DEFINITION	Hyaluronic acid synthetase modified protein.		
ACCESSION	E30971		
VERSION	E30971.1	GI:13017286	
KEYWORDS	JP 200004886-A/1.		
SOURCE	JP 200004886-A/1.		
ORGANISM	Mus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Naoki, I., Mamoru, Y. and Koji, K.		
TITLE	Hyaluronic acid synthetase modified protein		
JOURNAL	Patent: JP 200004886-A 1 11-JAN-2000;		
COMMENT	SEIKAGAKU KOGYO CO LTD OS Mus sp. (mouse) PN JP 200004886-A/1 PD 11-JAN-2000 PF 24-JUN-1998 JP 1998193788 PR PI NAOKI ITANO, MAMORU YOSHIDA, KOJI KIMATA PC C12N15/09, C12N9/00, C12N15/00 CC Strandedness: Double; CC Topology: Linear; FH Key FT CDS		
FEATURES	Location/Qualifiers		
source	1..2102		
ORIGIN	/organism="Mus sp." /mol_type="genomic DNA" /db_xref="taxon:10095"		
Query Match	62.6%;	Score 677.8;	DB 6;
Best Local Similarity	84.5%;	Pred. No. 1.8e-80;	
Matches 792;	Conservative	0; Mismatches 127;	Indels 18; Gaps 2
Qy	10	CAGACGGCGCCAGCCACTCCTGACGCCCGCTGCTCGGCGCTGCCCGGAGGTG	69
Db	55	CAGACATGCCAAGCCCTCAGAGCGCGGCTTGTCTGCTGTGGCTGTGCCAGGCGAGA	114
Qy	70	CTGACCATCGCTTCGCGCTGTCTATCTCTGGGCTCATGACCTGGGCTTACGCCCGCGG	129

	Matches	792;	Conservative	0;	Mismatches	127;	Indels	18;	Gaps	2;
QY	10	CAGACGCGCCCAAGCCCACTCTCTGCAGCCGCGCGCTGCTCCGCCCTGGCCCGAGGGTG	69							
Db	55	CAGACATGCGCAAGCCCTCAGAGGCAGCGGTTGCTGCTCTGGCCCTGGCCAGCGCAGCA	114							
QY	70	CTGACCATCGCTTCGCCCTGCTCATCTGGGCGCTCATGACCTGGGCGCTACGCCCGCGG	129							
Db	115	CTCAGCATCATCTTTGCCCCCTCATCTCTGGGCGCTCATGACCTGGGCGCTACGCCGAGGC	174							
QY	130	GTGCGGCTGGCGCTCCGATCGTACGGCGCTCTCTGGCGCTTCAGCGGCGCTTCCTT	189							
Db	175	GTTCCTCTGGCTTCAGATCGTATGACTCTCTGGCGCTTGGCGCTCTATGGGCATTCCTC	234							
QY	190	TCAGCGCACTGTGTGGCGCAGAGCCTCTTTCGCGTACTCTGAGACACCGCGCGGTGGCGCG	249							
Db	235	AGCGCACACCTAGTGTGGCACAGAGCCTCTTCGCTTACCTGTGAGCACCGAAGGTTGGCAGCG	294							
QY	250	GCGGCGG-----GGGCGCGCTGATGCAGCCACCGCGCGCGCAGTGTGGCG	294							
Db	295	GCTGCGCGCGCGCTCTTTGGCGAAGGGGCGCCCTGATGCGGCCACTGCACGCGAGCGTGGCA	354							
QY	295	CTGACCATCTCCGCTTACAGGAGGACCCCGCGTACCTGCGCCAGTGCCTGGCGCTCCGCC	354							
Db	355	CTCACCATCTCAGCTACACAGAGATCCCGCTTACTTGGCCAGTGTCTTGACCTTCGCG	414							
QY	355	CGGCGCTGTGTATACCGCGCGCGCGCTGCGCGTCTCTATGTGTGTGATGTGCAACCGC	414							
Db	415	CGCGCCTTGCTTACCCGCGCACAGAGGTTACGCGTGTCTATGTGTGTGACGCGCAACCGC	474							
QY	415	GCCGAGGACCTCTACATGTGTGCATCTTCGCGAGGTCTTCGCTGACGAGGACCCCGC	474							
Db	475	GCTGAGGATCTGTACATGTGTGAGCATGTTCGAGAGTCTTCGCCGATGAGGACCCCGC	534							
QY	475	ACGTACTGTGTGGACGGCAACTACACAGCCCTGGGAAACCGCGCGCG-----CGGGCGCG	531							
Db	535	ACTTATGTGTGGATGCGCACTACCATCAGCCCTGGGNAACAGCGAGGCTACGGCGCT	594							
QY	532	GTGGCGCCCGGAGCCTATCGGAGGTTGAGCGCGAGGATCTGTGGCGCGCTGGCAGTGAG	591							
Db	595	GTCCGTGAAGGTGCTTACCGGGAGTGTGAGCGGAGGACCCCGCGCGGTGTGGCGGTGGAG	654							
QY	592	GCGCTGTGTAGGACTCCGACGTGCTGCTGTGCGCAGCGCTGGGGCGCAACGCGAG	651							
Db	655	GCGCTGTGTAGAACACGACAGTGGTGTGCTGTGGCTCAGCGTTTGGGGCGCAACGTGAG	714							
QY	652	GTCACTGTACAGCGCTTCAAGGCGCTCCGAGATTTCGGTGGACTACGTGCGAGGTCTGTGAC	711							
Db	715	GTCACTGTACACAGCTTTCAGGGCACTGGCGACTCCGTGGACTACGTGCGAGTCTGTGAC	774							
QY	712	TCGGACACAAAGTTGTGACCCCATGCGACTCTGTGTGAGTCTGTGGGCTACTGAGCAGAGAC	771							
Db	775	TCAGACAAAGACTTAGACCCCATGGCACTGTGTGAGCTTGTGTGCGAGTGTGTGATGAAGAC	834							
QY	772	CCCGGGTACGGGCTGTGTGTGGGAGTGGCGGATCCTTAACCTCTGGAATCTCTGGGTC	831							
Db	835	CCCGGGTACGGGCTGTGTGTGAGGGGATGTGAGGATCCTTAACCTCTGGAATCTCTGGGTC	894							
QY	832	AGCTTCCTAAGCAGCCTTCGATCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGC	891							
Db	895	AGCTTCCTTGGAGCTTCTCGATCTGGGTAGCCTTCAATGTGGAAACGAGCTTGTGAGAGC	954							
QY	892	TACTTCACCTGTGTATCCCTGCATCAGCGGTCTCTTAG	928							
Db	955	TACTTCACCTGTGTGCTCTGCATCAGTGGTCTCTCTGG	991							

RESULT 14			
AC130783/c			
LOCUS	AC130783	19386 bp	DNA linear
DEFINITION	Pan troglodytes clone CH251-426A12,	WORKING DRAFT SEQUENCE,	8
	ordered pieces.		
ACCESSION	AC130783		

VERSION	REFERENCE	TITLE	REFERENCE	REFERENCE	REFERENCE
KEYWORDS	AUTHORS	JOURNAL	AUTHORS	AUTHORS	AUTHORS
SOURCE			TITLE	TITLE	TITLE
ORGANISM			JOURNAL	JOURNAL	JOURNAL
					COMMENT

AC130783.2 GI:25100968
 HTG: HTGS_PHASE2; HTGS_DRAFT.
 Pan troglodytes (chimpanzee)
 Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1 (bases 1 to 139986)
 Akter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
 Benjamin,B., Blakestei,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
 Carriaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
 Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
 Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
 Marquies,E.H., Mastello,C., Maskeri,B., McDowell,J.,
 Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
 Reddi-Dugue,N., Schander,K., Schueler,M.G., Sison,C.,
 Stantrypop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
 Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 139986)
 Green,E.D.
 Direct Submission
 Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 139986)
 Green,E.D.
 Direct Submission
 Submitted (19-NOV-2002) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 On Nov 19, 2002 this sequence version replaced gi:22218453.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@nhgri.nih.gov
 ----- Project information
 Center project name: dxd
 Center clone name: 426A12

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.930319
Consensus quality: 191764 bases at least Q40
Consensus quality: 192510 bases at least Q30
Consensus quality: 192998 bases at least Q20
Insert size: 184000; agarose-fp
Insert size: 193286; sum-of-contigs
Quality coverage: 11.49x in Q20 bases; agarose-fp
Quality coverage: 10.94x in Q20 bases; sum-of-contigs

```

```

* -----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*      1      11512: contig of 11512 bp in length

```

```
* 11513 11612: gap of unknown length
* 5988: contig of 47876 bp in length
* 11613 5988: gap of unknown length
* 59889 103266: contig of 43678 bp in length
* 103267 127885: contig of 24519 bp in length
* 127886 127985: gap of unknown length
* 127986 129692: contig of 1707 bp in length
* 129692 129792: gap of unknown length
* 129792 131077: contig of 1285 bp in length
* 131078 174663: contig of 43486 bp in length
* 174664 174764: gap of unknown length
* 174764 193986: contig of 19223 bp in length.
FEATURES
    source
        1..193986
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="CH251-426A12"
            /clone_lib="CH251"
            1..11512
                /note="assembly_fragment"
                clone_end:5p6
                vector_side:left
            1..76
                /note="clone overlaps with GenBank Accession Number
                AC132069 clone CH251-132A6 (center project name dxc)"
                11613..59488
                    /note="assembly_fragment"
                    59589..103266
                        /note="assembly_fragment"
                        103367..127885
                            /note="assembly_fragment"
                            127986..129692
                                /note="assembly_fragment"
                                129793..131077
                                    /note="assembly_fragment"
                                    131178..174663
                                        /note="assembly_fragment"
                                        174764..193986
                                            /note="assembly_fragment"
                                            clone_end:T7
                                            vector_side:right
ORIGIN
Query Match 61.8%; Score 669; DB 2; Length 193986;
Best Local Similarity 97.2%; Pred. No. 6.6e-80;
Matches 697; Conservative 0; Mismatches 5; Indels 15; Gaps 1;
Qy 7 CAGCAGGACGGCCGACAGCCACTCTCTGACGCCCGCGCTCTCGGCTTGGCCGGAGG 132769
Db 132828 CAGCAGGACGGCCGACAGCCACTCTCTGACGCCCGCGCTCTCGGCTTGGCCGGAGG 132769
Qy 67 GTGCTGACCATCGCTTGGCTTGCATCTCTGGCTCATGACCTGAGCTAGCGCCGCGCC 126
Db 132768 GTGCTGACCATCGCTTGGCTTGCATCTCTGGCTCATGACCTGAGCTAGCGCCGCGCC 132709
Qy 127 GGGGTGCGCTGGCTTGCATCTGCTACGGCTTCTGGCTTCTGGCTTCTAGCGGCTTTC 186
Db 132708 GGGGTGCGCTGGCTTGCATCTGCTACGGCTTCTGGCTTCTGGCTTCTAGCGGCTTTC 132649
Qy 187 CTTTCAGCGCACTGTGGTGGCGAGAGCTTCTTTCGGGTACTGTGAGACACCGCGGGTGGCG 246
Db 132648 CTTTCAGCGCACTGTGGTGGCGAGAGCTTCTTTCGGGTACTGTGAGACACCGCGGGTGGCG 132589
Qy 247 -----GCGCGCGCGCGCGCGCGCGCGCTTGGATGACGACCGCGCGAGTGTG 291
Db 132588 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGGATGACGACCGCGCGAGTGTG 132529
Qy 292 GCGCTGACCATCTCCGCTTACAGGAGACCCCGCTACTGTCGCGCAGTGCCTGGCGTCC 351
Db 132528 GCGCTGACCATCTCCGCTTACAGGAGACCCCGCTACTGTCGCGCAGTGCCTGGCGTCC 132469

Qy 352 GCCCGCGCTGTGTACCGCGCGCGCTGCGCTCTCATGTGGTGGTGGCAAC 411
Db 132468 GCCCGCGCTGTGTGTACCGCGCGCGCTGCGCTCTCATGTGGTGGTGGCAAC 132409
Qy 412 CGCGCCGAGGACCTCTACATGTTCGACATGTTCGCGAGAGTCTTCGCTGACGAGGACCCC 471
Db 132408 CGCGCCGAGGACCTCTACATGTTCGACATGTTCGCGAGAGTCTTCGCTGACGAGGACCCC 132349
Qy 472 GCCACGTACTGTGGGACGGCAACTACACAGCCCTGGGAAACCCGCGCGCGCGCGCGCG 531
Db 132348 GCCACGTACTGTGGGACGGCAACTACACAGCCCTGGGAAACCCGCGCGCGCGCGCGCG 132289
Qy 532 GTGGCGCGCGGAGCTATCGGAGGTGAGCGGAGGATCTCTGGCGGCTGGCAGTGTGAG 591
Db 132288 GTGGCGCGCGGAGCTATCGGAGGTGAGCGGAGGATCTCTGGCGGCTGGCAGTGTGAG 132229
Qy 592 GCGCTGTGTGAGGACTCGCAGGTGCTGTGGTGGCGCAGCGCTGGGCGGCAAGCGCGAG 651
Db 132228 GCGCTGTGTGAGGACTCGCAGGTGCTGTGGTGGCGCAGCGCTGGGCGGCAAGCGCGAG 132169
Qy 652 GTCATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAGGTCTGT 708
Db 132168 GTCATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAGGTGAGT 132112

RESULT 15
AC137058/c
LOCUS
DEFINITION Papio anubis clone RP41-126M5, WORKING DRAFT SEQUENCE, 10 ordered
pieces.
ACCESSION AC137058
VERSION AC137058.3 GI:28460766
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Papio anubis (olive baboon)
ORGANISM Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
REFERENCE 1 (bases 1 to 185623)
AUTHORS Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.I., McDowell,J.,
Margulies,E.H., Masiello,C., Maskeri,B., Prasad,A.,
Paguigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
REFERENCE 2 (bases 1 to 185623)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 185623)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Feb 21, 2003 this sequence version replaced gi:27476124.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: dtj
Center clone name: 126M05

The sequence data in this record represents an 'enhanced'
```


version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 183092 bases at least Q40
Consensus quality: 183851 bases at least Q30
Consensus quality: 184434 bases at least Q20
Insert size: 160000; agarose-fp
Insert size: 184723; sum-of-contigs
Quality coverage: 14.13x in Q20 bases; agarose-fp
Quality coverage: 12.24x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 9984: contig of 9984 bp in length
* 9985 10084: gap of unknown length
* 10085 49805: contig of 3721 bp in length
* 49806 49905: gap of unknown length
* 49906 66936: contig of 17031 bp in length
* 66937 67036: gap of unknown length
* 67037 83604: contig of 16568 bp in length
* 83605 83704: gap of unknown length
* 83705 108246: contig of 24542 bp in length
* 108247 108346: gap of unknown length
* 108347 112021: contig of 3675 bp in length
* 112022 112121: gap of unknown length
* 112122 134817: contig of 22696 bp in length
* 134818 134918: gap of unknown length
* 134918 147290: contig of 12373 bp in length
* 147291 147390: gap of unknown length
* 147391 185337: contig of 37947 bp in length
* 185338 185437: gap of unknown length
* 185438 185623: contig of 186 bp in length.

FEATURES

Location/Qualifiers
1..185623
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="RP41-126W5"
/clone_lib="RP41"

1..15263
/note="clone overlaps with GenBank Accession Number AC130273 clone RP41-30706 (center project name dsu)"

1..9984

/note="assembly_fragment
clone_end:77
vector_side:left"

10085..49805
/note="assembly_fragment"
49906..66936
/note="assembly_fragment"
67037..83604
/note="assembly_fragment"
83705..108246
/note="assembly_fragment"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature 108347..112021
/note="assembly_fragment"
misc_feature 112122..134817
/note="assembly_fragment"
misc_feature 134918..147290
/note="assembly_fragment"
misc_feature 147391..185337
/note="assembly_fragment"
misc_feature 185438..185623
/note="assembly_fragment
clone_end:SP6
vector_side:right"

ORIGIN

Query Match 58.7%; Score 635.4; DB 2; Length 185623;
Best Local Similarity 94.3%; Pred. No. 1.8e-75;
Matches 676; Conservative 0; Mismatches 26; Indels 15; Gaps 1;

QY 7 CAGCAGGACGCGCCAAAGCCCACTCTCTGACGCGCGCTGCTCGGCTTGGCCGAGG 66
Db 138595 CAGCGGACACGCGCCAAAGCCCACTCTCTGACGCGCGCTGCTCGGCTTGGCCGAGG 138536
QY 67 GTGCTGACCATCGCTTGGCCCTGCTCATCTCTGGGCTCATGACTGGCGCTTACGCCGC 126
Db 138535 GTGCTGACCATCGCTTGGCCCTGCTCATCTCTGGGCTCATGACTGGCGCTTACGCCGC 138476
QY 127 GGGGTGCGCGTGGGCTCCGATCGCTACGGCTCTCGGCTTTCGGCTCTACGGGGCTTC 186
Db 138475 GGGGTGCGCGTGGGCTCCGATCGCTACGGCTCTCTCGGCTTTCGGCTCTACGGGGCTTC 138416
QY 187 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTTCGGCTACCTTGGAGCACCGCGGCTGCGG 246
Db 138415 CTCTCGCGCGACCTGTTGGCGCAGAGCTCTTTCGGCTACCTTGGAGCACCGCGGCTGCGG 138356
QY 247 CCGCGCGCGG-----CGGGGCGCGTGGATGAGCCACCGCGCGCTGTG 291
Db 138355 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTGAGTACCGCGCGCGCTGTG 138296
QY 292 GCGCTGACCATCTCCGCTTACAGGAGGACCGCGCTACCTCGCGCTGCTGCTGCGCTGCC 351
Db 138295 GCGCTGACCATCTCCGCTTACAGGAGGACCGCGCTGCTGCTGCGCTGCCGCTGCC 138236
QY 352 GCCCGCGCGCTGCTGTATCCCGCGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 411
Db 138235 GCCCGCGCGCTGCTGTATCCCGCGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 138176
QY 412 CGCGCGGAGGACCTTACATGCTGACATGTTTCGGAGGCTTTCGCTGACGAGGACCGCC 471
Db 138175 CGCGCGGAGGACCTTACATGCTGACATGTTTCGGAGGCTTTCGCTGACGAGGACCGCC 138116
QY 472 GCCACGTACGTGTGGGACGGCAACTACCCAGACCGCTGGGAAACCGCGCGCGCGCGCG 531
Db 138115 GCCACGTACGTGTGGGACGGCAACTACCCAGACCGCTGGGAAACCGCGCGCGCGCGCG 138056
QY 532 GTGGCGCGCGGAGCTTATCGGGAGGTGAGCGGAGGATCTTGGCGCGCTGCGCTGCGCTGAG 591
Db 138055 GTGGCGCGCGGAGCTTATCGGGAGGTGAGCGGAGGATCTTGGCGCGCTGCGCTGCGCTGAG 137996
QY 592 GCGCTGTGAGGACTCGCAGGTGCTGCTGCGCGCAGCGCTGCGGCGCGCGCGCGCGAG 651
Db 137995 GCGCTGTGAGGACTCGCAGGTGCTGCTGCGCGCAGCGCTGCGGCGCGCGCGCGAG 137936
QY 652 GTCATGTACAGCGCTTCAAGCGCTCGGAGATTCGGTGGACTAGTCTGAGTCTGT 708
Db 137935 GTCATGTACAGCGCTTCAAGCGCTCGGAGACTCGGTTGGACTAGTCTGAGTCTGT 137879

Search completed: March 13, 2005, 08:01:32
Job time : 5118.08 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 01:47:21 ; Search time 703.833 Seconds
(without alignments)
9108.809 Million cell updates/sec

Title: US-10-672-399-5

Perfect score: 1083

Sequence: 1 atgagacagcagcgccgccccccccc...tccaagtcgtactccgtga 1083

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	933	85.1	1737	10 ADL13691	Adl13691 Osteoarth
2	933	85.1	2088	10 ADL13690	Adl13690 Osteoarth
3	933	85.1	2116	2 AAT99541	Aat99541 Human hya
4	931.4	85.0	1737	10 ADL13692	Adl13692 Osteoarth
5	931.4	85.0	2087	10 ADL13694	Adl13694 Osteoarth
6	911	84.1	2117	2 AAT96713	Aat96713 Human hya
7	697.2	64.4	23122	8 ADL13693	Adl13693 Osteoarth
8	677.8	62.6	1752	8 ABZ76734	Abz76734 Mouse hya
9	677.8	62.6	1752	10 AAD59442	Aad59442 Mouse hya
10	677.8	62.6	2102	2 AAT91855	Aat91855 Mouse hya
11	677.8	62.6	2102	2 AAZ10862	Aaz10862 Hyaluro
12	677.8	62.6	2102	3 AAZ88199	Aaz88199 Mouse hya
13	677.8	62.6	2102	3 AAA33987	Aaa33987 Murine HA
14	422	39.0	662	6 ABQ27384	Abq27384 Oligonuc
15	422	39.0	662	6 ABQ27385	Abq27385 Oligonuc
16	393.8	36.4	662	6 ABQ27382	Abq27382 Oligonuc
17	393.8	36.4	662	6 ABQ27383	Abq27383 Oligonuc
18	333.2	31.3	490	9 ACH39998	Ach39998 Human foe
19	290.8	26.9	1783	10 ADC49212	Adc49212 Rabbit hy
20	273	25.2	1662	8 ACC51023	Acc51023 Human bla

21	273	25.2	1662	8 ABX76415	Abx76415 Lung canc
22	273	25.2	1662	11 ADN39051	Adn39051 Cancer/an
23	273	25.2	4049	8 ABX76214	Abx76214 Lung canc
24	273	25.2	4049	12 ADN05115	Adn05115 Antipgori
25	273	25.2	4220	12 ADJ74905	Adj74905 Marker ge
26	250	23.1	1767	10 ADD93927	Add93927 Xenopus l
27	236.4	21.8	1665	2 AAV18822	Aav18822 Murine hy
28	236.4	21.8	1665	3 AAZ88201	Aaz88201 Mouse hya
29	236.4	21.8	1665	8 ABZ76736	Abz76736 Mouse hya
30	236.4	21.8	1665	10 AAD59444	Aad59444 Mouse hya
31	236.4	21.8	5919	12 ADJ75814	Adj75814 Marker ge
32	207.4	19.2	1229	12 ADJ74906	Adj74906 Marker ge
33	179.4	16.6	1653	6 ABL41013	Ab141013 Murine hy
34	179.4	16.6	1659	8 ABZ76735	Abz76735 Mouse hya
35	179.4	16.6	1659	10 AAD59443	Aad59443 Mouse hya
36	179.4	16.6	2947	2 AAV18821	Aav18821 Murine hy
37	179.4	16.6	2948	6 ABI98697	Abi98697 Mouse isc
38	179.4	16.6	4194	9 ADA49685	Ada49685 DNA encod
39	179.4	16.6	4194	10 ADC59298	Adc59298 Mouse HAS
40	165.8	15.3	2890	3 AAZ88200	Aaz88200 Mouse hya
41	153.6	14.2	1659	10 ADL13696	Adl13696 Osteoarth
42	153.6	14.2	3003	6 ABK84373	Abk84373 Human CDN
43	153.6	14.2	3003	9 ADA49683	Ada49683 DNA encod
44	153.6	14.2	3003	10 ADC59296	Adc59296 Human HAS
45	153.6	14.2	3003	10 ADD12718	Add12718 Human CDN

ALIGNMENTS

RESULT 1

ADL13691

ID ADL13691 standard; DNA; 1737 BP.

XX AC ADL13691;

XX DT 06-MAY-2004 (first entry)

XX DE Osteoarthritis-associated polymorphic nucleotide #223.

XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX joint space narrowing; osteophyte development; joint pain;
XX osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX (INCY-) INCYTE GENOMICS INC.
XX Jones KA, Schafer A;
XX WPI; 2003-559141/52.
XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.
XX Disclosure; SEQ ID NO 223; 297pp; English.
XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space

narrowing and/or osteophyte development and/or joint pain that is associated with a disease, preferably osteoarthritis. The cell line and the non-human animal are useful for screening for an agent for diagnosing an individual having susceptibility to joint space narrowing and/or osteophyte development and/or joint pain. This sequence corresponds to the polypeptide encoding a protein listed in the specification. (Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences/](http://www.wipo.int/pub/published_pct_sequences/)).

Sequence 1737 BP; 232 A; 571 C; 603 G; 331 T; 0 U; 0 Other; 0

Query Match 86.1%; Score 933; DB 10; Length 1737;

Best Local Similarity 90.6%; Pred. No. 2e-151;

Matches	1048;	Conservative	0;	Mismatches	35;	Indels	74;	Gaps	2;
---------	-------	--------------	----	------------	-----	--------	-----	------	----

Qy	1	ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGACCGCGCGCTGCTCTCGCGCTGCGC	60
Db	1	ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGACCGCGCGCTGCTCTCGCGCTGCGC	60
Qy	61	CGGAGGTGTGACACATCGCTTTCGCCCTGTCTCATCTTGGGCGCTCATGACCTTGGGCGCTTAC	120
Db	61	CGGAGGTGTGACCAATCGCTTTCGCCCTGTCTCATCTTGGGCGCTCATGACCTTGGGCGCTTAC	120
Qy	121	GCGCCCGGGTGCCTCGCTGGCTTCGATCGCTACGCGCTCTCTGGGCTTTCGGGCTCTTACGGG	180
Db	121	GCGCCCGGGTGCCTCGCTGGCTTCGATCGCTACGCGCTCTCTGGGCTTTCGGGCTCTTACGGG	180
Qy	181	GCCTTCTTTTCAGCGCACCTTGGTGGCGCAGAGCCCTTTCGGGTACCTGGAGCACCGGCGG	240
Db	181	GCCTTCTTTTCAGCGCACCTTGGTGGCGCAGAGCCCTTTCGGGTACCTGGAGCACCGGCGG	240
Qy	241	GTGGCGGCGCGCGCGGGGCGCTGTGATGCAGCACCGCGCGCAGTGTGTGGCGTGA	300
Db	241	GTGGCGGCGCGCGCGGGGCGCTGTGATGCAGCACCGCGCGCAGTGTGTGGCGTGA	300
Qy	301	ATCTCCGCTTACAGAGGACACCCCGCGTACCTTGGGCCAGTGTGTGGCGTGTGGCGCGC	360
Db	301	ATCTCCGCTTACAGAGGACACCCCGCGTACCTTGGGCCAGTGTGTGGCGTGTGGCGCGC	360
Qy	361	CTGCTGTACCCGCGCGCGCGCTGCGCGTCTCTATGGTGTGTGGTGGCAACCGCGCGCGAG	420
Db	361	CTGCTGTACCCGCGCGCGCGCTGCGCGTCTCTATGGTGTGTGGTGGCAACCGCGCGCGAG	420
Qy	421	GACCTTCACTGGTTCGACATGTTCCGCGAGTCTTTCGCTGACGAGGACCCCGCCACGTAC	480
Db	421	GACCTTCACTGGTTCGACATGTTCCGCGAGTCTTTCGCTGACGAGGACCCCGCCACGTAC	480
Qy	481	GTGTGGGACGCGCACTACCAACGCGCTTGGGACCCGCGCGCGGGCGGGCGGTGGGCGCC	540
Db	481	GTGTGGGACGCGCACTACCAACGCGCTTGGGACCCGCGCGCGGGCGGGCGGTGGGCGCC	540
Qy	541	GGAGCTTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGGCAGTGTGAGGCGCTGTGTG	600
Db	541	GGAGCTTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGGCAGTGTGAGGCGCTGTGTG	600
Qy	601	AGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGCGCAACGCGAGTCAATGTAC	660
Db	601	AGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGCGCAACGCGAGTCAATGTAC	660
Qy	661	ACAGCTTTCAGGGCGCTCGGAGATTCGGTGGACTACGTGCACTGTGTGACTTCGACACACA	720
Db	661	ACAGCTTTCAGGGCGCTCGGAGATTCGGTGGACTACGTGCACTGTGTGACTTCGACACACA	720
Qy	721	AGGTTGGACCCCATGGCACTGCTGGAGTCTGTGGCGGTACTTGGACGAGGACCCCGGGTAC	780
Db	721	AGGTTGGACCCCATGGCACTGCTGGAGTCTGTGGCGGTACTTGGACGAGGACCCCGGGTAC	780
Qy	781	GGGCGTGTGTGGGACGTGCGGATCTCTTAACCTCTGGACTCTCTGGGTGAGCTTCCTA	840
Db	781	GGGCGTGTGTGGGACGTGCGGATCTCTTAACCTCTGGACTCTCTGGGTGAGCTTCCTA	840
Qy	841	AGCAGCTTCGATACTGTGTTAGCTTCAATGTGGAGCGGGCTTGTTCAGAGCTACTTCCAC	900

CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences/.

XX Sequence 2088 BP; 302 A; 652 C; 711 G; 423 T; 0 U; 0 Other;

Query Match 86.1%; Score 933; DB 10; Length 2088;
Best Local Similarity 90.6%; Pred. No. 2e-151;
Matches 1048; Conservative 0; Mismatches 35; Indels 74; Gaps 2;

```
QY 1 ATGACAGCAGGAGCGCCCAAGCCCACTCTGTCAGCCCGCCGCTGCTCCGGCTGGCC 60
Dd 36 ATGACAGCAGGAGCGCCCAAGCCCACTCTGTCAGCCCGCCGCTGCTCCGGCTGGCC 95
QY 61 CGAGGGGTGCTGACCATCGCCTTCGCCCTTCGCTGTCATCTGGGCTCATGACCTGGGCTTAC 120
Dd 96 CGAGGGGTGCTGACCATCGCCTTCGCCCTTCGCTGTCATCTGGGCTCATGACCTGGGCTTAC 155
QY 121 GCCGCGGGGTGCTGCTGCTCCGATCGCTACGCGCTCTCGGCTTCGGCTCTTACGGG 180
Dd 156 GCCGCGGGGTGCTGCTGCTCCGATCGCTACGCGCTCTCGGCTTCGGCTCTTACGGG 215
QY 181 GCCTTCTTTCAGCGCACCTGCTGGCGCAGAGCCTCTTCGCTACCTGGAGCACCGCGG 240
Dd 216 GCCTTCTTTCAGCGCACCTGCTGGCGCAGAGCCTCTTCGCTACCTGGAGCACCGCGG 275
QY 241 GTGGCGGGCGCGCGGGCGCTGGATGACGACCGCGCGAGTGTGGCGCTGACC 300
Dd 276 GTGGCGGGCGCGCGGGCGCTGGATGACGACCGCGCGAGTGTGGCGCTGACC 335
QY 301 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGCTCGCGCTCGCGCGCC 360
Dd 336 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGCTCGCGCTCGCGCGCC 395
QY 361 CTGCTGTACCGCGCGCGGGTGGCTGCTCTCATGCTGGTGGTGGTGGTGGTGGTGGTGG 420
Dd 396 CTGCTGTACCGCGCGCGGGTGGCTGCTCTCATGCTGGTGGTGGTGGTGGTGGTGGTGG 455
QY 421 GACCTCTACATGCTGCAATGCTTCGCGAGAGTCTTCGCTGACGAGAGCCCGCCACGTAC 480
Dd 456 GACCTCTACATGCTGCAATGCTTCGCGAGAGTCTTCGCTGACGAGAGCCCGCCACGTAC 515
QY 481 GTGTGGGACGGCAATACACACAGCCCTGGGAAACCCCGCGCGCGCGCGCGCGCGCGCG 540
Dd 516 GTGTGGGACGGCAATACACACAGCCCTGGGAAACCCCGCGCGCGCGCGCGCGCGCGCG 575
QY 541 CGAGCTATCGGAGGTGGAGCGAGGATCTGCGCGCGCTGGCAGTGGAGCGCTGGTG 600
Dd 576 GGAGCTATCGGAGGTGGAGCGAGGATCTGCGCGCGCTGGCAGTGGAGCGCTGGTG 635
QY 601 AGGACTCGCAGGTGCTGCTGCGCGCAGCGCTGGGCGGCAAGCGGAGGTCTATGTAC 660
Dd 636 AGGACTCGCAGGTGCTGCTGCGCGCAGCGCTGGGCGGCAAGCGGAGGTCTATGTAC 695
QY 661 ACAGCCTTCAAGCGCTCGGAGATTGGTGGTACGTCAGGTCGTGACTCGGACACA 720
Dd 696 ACAGCCTTCAAGCGCTCGGAGATTGGTGGTACGTCAGGTCGTGACTCGGACACA 755
QY 721 AGTTGGACCCCATGCGACTGCTGAGCTGCTGGCGGTACTGGACGAGGACCCCGGTA 780
Dd 756 AGTTGGACCCCATGCGACTGCTGAGCTGCTGGCGGTACTGGACGAGGACCCCGGTA 815
QY 781 GGGGCTGTGGTGGGAGCTGGGATCTTAACCTCTGGACTCTCGGCTGAGCTTCTTA 840
Dd 816 GGGGCTGTGGTGGGAGCTGGGATCTTAACCTCTGGACTCTCGGCTGAGCTTCTTA 875
QY 841 AGCAGCTTCGATCTGGGTAGCCTTCAATGTGGAGCGGGCTGTGACAGCTTCTTCCAC 900
Dd 876 AGCAGCTTCGATCTGGGTAGCCTTCAATGTGGAGCGGGCTGTGACAGCTTCTTCCAC 935
QY 901 TGTGTATCTGATCAGCGGCTCTTAG----- 928
Dd 936 TGTGTATCTGATCAGCGGCTCTTAGCCCTATATAGGAATAACCTCTTGCAGCAGTTT 995
```

```
QY 929 -----AATCTGCCAGGCCCCCAGGAGCAGCGATGATGCC 965
Dd 996 CTTGAGGCTGGTACACCAAGAGTCTCTGGTACCCACTGTACTTTTGGGATGACCGG 1055
QY 966 CTCATTCTCGCCC-----CGTCAGGTACACCTCCAGGTCCCGC 1006
Dd 1056 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTACACCTCCAGGTCCCGC 1115
QY 1007 TGCTACTCAGAGAGCGCCCTGCTCTCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1066
Dd 1116 TGCTACTCAGAGAGCGCCCTGCTCTCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1175
QY 1067 AAGTCGTACTTCCGTGA 1083
Dd 1176 AAGTCGTACTTCCGTGA 1192
```

RESULT 3

AAT99541

ID AAT99541 standard; cDNA; 2116 BP.

XX

AC AAT99541;

XX

DT 21-MAY-1998 (first entry)

XX

XX Human hyaluronan synthase cDNA clone 30C.

DE

KW Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
wound healing; vulnery; tissue repair; scar; keloid; therapy; ds.

OS

XX Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 36..1172

FT polyA_signal /*tag= a

FT /*tag= b

XX

PN W09740174-A1.

XX

PD 30-OCT-1997.

XX

PF 17-APR-1997; 97WO-US006350.

XX

PR 22-APR-1996; 96US-00635552.

XX

PA (LEUK-) LEUKOSITE INC.

XX

PI Briekin MJ;

XX

DR WPI, 1997-549359/50.

XX

DR P-PSDB; AAW26765.

XX

PT Human hyaluronan synthase - useful for recombinant production of
hypertrophic scar and keloid formation.

XX

PS Claim 3; Page 36-38; 58pp; English.

XX

CC cDNA clone 30C includes a coding region for hyaluronan synthase (HAS)
(see AAW26765), an enzyme involved in the synthesis of hyaluronan
(hyaluronic acid) and which has the ability to confer cell adhesion by
the lymphocyte receptor CD44. Clone 30C was isolated using an expression
cloning system developed to isolate cDNA clones that encode proteins that
confer adhesion of the murine T cell lymphoma TK1. A human mesenteric
lymph node expression library was constructed that, upon transfection
into CHO/P cells, yielded clone 30C that mediated rosetting of TK1 cells
to some of the transfectants. The isolated clone can be utilised in a
claimed method for producing HAS in host cells. Such host cells are used
in a claimed method for the production of hyaluronan. Hyaluronan is
useful for wound healing and tissue repair, and can reduce or prevent
hypertrophic scars and keloid formation. It is also used in eye surgery

QY 61 CGGAGGGTGTGACCATCGCCTTTCGCCCTGCTCATCTCGGGCTCATGACCTGGGCGCTAC 120
DB |||||||
QY 96 CGGAGGGTGTGACCATCGCCTTTCGCCCTGCTCATCTCGGGCTCATGACCTGGGCGCTAC 155
DB |||||||
QY 121 GCGCGCGGGGTGCGCTCGCCTCCGATCGCTACGGCCTCTGGCCTCTGGCCTCTACGGG 180
DB |||||||
QY 156 GCGCGCGGGGTGCGCTCGCCTCCGATCGCTACGGCCTCTGGCCTCTGGCCTCTACGGG 215
DB |||||||
QY 181 GCCTTCCTTTAGCGCAGCTGTGTGGCGAGAGCCTCTTCGCGTACTCTGGAGCACCGGGGG 240
DB |||||||
QY 216 GCCTTCCTTTAGCGCAGCTGTGTGGCGAGAGCCTCTTCGCGTACTCTGGAGCACCGGGGG 275
DB |||||||
QY 241 GTGGCGGGCGGGCGGGCGGGCGCTGTGATGACGACCGCGCGAGTGTGCGCTGACC 300
DB |||||||
QY 276 GTGGCGGGCGGGCGGGCGGGCGCTGTGATGACGACCGCGCGAGTGTGCGCTGACC 335
DB |||||||
QY 301 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGTCTGGCGTCCGCGCGGCC 360
DB |||||||
QY 336 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGTCTGGCGTCCGCGCGGCC 395
DB |||||||
QY 361 CTGCTGTACCGCGCGGGGTGCGCTCTCATGTGTGTGGATGGCAACCGCGCGGAG 420
DB |||||||
QY 396 CTGCTGTACCGCGCGGGGTGCGCTCTCATGTGTGTGGATGGCAACCGCGCGGAG 455
DB |||||||
QY 421 GACCTCTACATGCTGACATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCCACGTAC 480
DB |||||||
QY 456 GACCTCTACATGCTGACATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCCACGTAC 515
DB |||||||
QY 481 GTGTGGGACGGCAATACACAGCCCTTGGGAACCCCGCGGGCGGGCGCGTGGCGGCC 540
DB |||||||
QY 516 GTGTGGGACGGCAATACACAGCCCTTGGGAACCCCGCGGGCGGGCGCGTGGCGGCC 575
DB |||||||
QY 541 GAGGCTTATCGGAGGTGAGCGGAGGATCTTGGCGGGCTGGCAGTGGAGGCGCTGTGTG 600
DB |||||||
QY 576 GAGGCTTATCGGAGGTGAGCGGAGGATCTTGGCGGGCTGGCAGTGGAGGCGCTGTGTG 635
DB |||||||
QY 601 AGGACTCCGAGTGTGTGCGTGGCGAGCGCTGGGGCGGCGAGAGTGTATGTAC 660
DB |||||||
QY 636 AGGACTCCGAGTGTGTGCGTGGCGAGCGCTGGGGCGGCGAGAGTGTATGTAC 695
DB |||||||
QY 661 ACAGCCTTCAAGCGCTCGGAGATTTCGGTGGACTACGTGCGAGGTCTGTGACTCGGACACA 720
DB |||||||
QY 696 ACAGCCTTCAAGCGCTCGGAGATTTCGGTGGACTACGTGCGAGGTCTGTGACTCGGACACA 755
DB |||||||
QY 721 AGTGTGAGCCCATGCGACTGTGAGCTCTGTGGGGTACTGGAGGAGACCCCGGGTA 780
DB |||||||
QY 756 AGTGTGAGCCCATGCGACTGTGAGCTCTGTGGGGTACTGGAGGAGACCCCGGGTA 815
DB |||||||
QY 781 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTGGAATCTGGAATCTGGAATCTGGAATCTGGA 840
DB |||||||
QY 816 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTGGAATCTGGAATCTGGAATCTGGAATCTGGA 875
DB |||||||
QY 841 AGCAGCTGCGATCTGCGGTAGCCTTCAATGTGAGCGGGCTTGTGAGAGTACTTCCAC 900
DB |||||||
QY 876 AGCAGCTGCGATCTGCGGTAGCCTTCAATGTGAGCGGGCTTGTGAGAGTACTTCCAC 935
DB |||||||
QY 901 TGTGTATCTGATCAGCGGTCTCTAG- 928
DB |||||||
QY 936 TGTGTATCTGATCAGCGGTCTCTAGCGGCTATATAGGAATAACCTCTTGCAGCAGTTT 995
DB |||||||
QY 929 -----AATCTGCCAGGCCCGGAGGACGCGATGATGCC 965
DB |||||||
QY 996 CTTGAGGCTGGTACAAACAGAGTTCTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1055
DB |||||||
QY 966 CTCATTCCTCGCCCC-----CGTGCAGGTACACCTCCAGGTCCCGC 1006
DB |||||||
QY 1056 CACCTCAACCAACCGATGCTCAGCATGGTGTATGTACCAAGTACACCTCCAGGTCCCGC 1115
DB |||||||
QY 1007 TGCTACTCAGAGACGCCCTCTGCTCTCTGCGGTGGCTGAGCGAGCAGACAGCTGTGTC 1066
DB |||||||
QY 1116 TGCTACTCAGAGACGCCCTCTGCTCTCTGCGGTGGCTGAGCCAGCAGACAGCTGTGTC 1175
DB |||||||
QY 1067 AAGTCGTACTTCCGTGA 1083

DB 1176 AAGTCGTACTTCCGTGA 1192
|||||
RESULT 6
AAT96713
ID AAT96713 standard; DNA; 2117 BP.
XX AAT96713;
AC AAT96713;
XX
DT 22-APR-1998 (first entry)
XX
DE Human hyaluronate synthetase coding sequence.
XX
KW Hyaluronate synthetase; human; hyaluronic acid; drug preparation;
cosmetic preparation; gene therapy; carcinogenesis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 149..1780
FT /*tag= a
XX
PN WO9738113-A1.
XX
PD 16-OCT-1997.
XX
PF 31-MAR-1997; 97WO-JP001111.
XX
PR 05-APR-1996; 96JP-00084326.
XX
PR 30-APR-1996; 96JP-00109663.
XX
PA (SEBK) SEIKAGAKU CORP.
XX
PI Itano N, Kimata K;
XX
XX WPI; 1997-512726/47.
DR P-PSDB; AAW36503.
XX
XX DNA encoding human hyaluronate synthetase - for industrial scale
production of hyaluronic acid used in generating anti-carcinogenic drugs
or for cosmetics.
XX
PS Claim 3; Page 23-27; 35pp; Japanese.
XX
XX This sequence encodes a human hyaluronate synthetase, and is the coding
sequence of the invention. The encoded enzyme is useful for industrial
scale production of hyaluronic acid for use in the preparation of drugs
and cosmetics. The drugs can also be used in compositions for the
treatment of disorders involving the lowering of hyaluronic acid
production. The peptides may be used for the preparation of antibodies
recognising hyaluronate synthetase, e.g. for diagnostic purposes.
XX
XX Antisense DNA or RNA corresponding to the DNA can be used in gene therapy
treatment of carcinogenesis
XX
SQ Sequence 2117 BP; 306 A; 673 C; 703 G; 435 T; 0 U; 0 Other;
Query Match 84.1%; Score 911; DB 2; Length 2117;
Best Local Similarity 89.7%; Pred. No. 1.2e-147;
Matches 1032; Conservative 0; Mismatches 45; Indels 74; Gaps 2;
QY 7 CAGCAGAGCGGCGCCAGCCACTCTCTGACCGCGCGCTGCTCGGCTGGCCCGGAGG 66
DB |||||||
QY 50 CAGCAGAGCGGCGCCAGCCACTCTCTGACCGCGCGCTGCTCGGCTGGCCCGGAGG 109
DB |||||||
QY 67 GTGTGACCATCGCCTTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTACGCGGCC 126
DB |||||||
QY 110 GTGTGACCATCGCCTTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTACGCGGCC 169
DB |||||||
QY 127 GGGGTGCGCTGGCCTCCGATCGCTACGGCTCTGCGCTTGGCCTTACGGGGCCTTC 186
DB |||||||
QY 170 GGGGTGCGCTGGCCTCCGATCGCTACGGCTCTGCGCTTGGGCTTACGGGGCCTTC 229
DB |||||||


```
QY 187 CTTTCAGCGCACCTGGTGGCGAGAGCCTCTTTCGGCTACCTGGAGACACGGCGGTGGCG 246
Db 230 CTTTCAGCGCACCTGGTGGCGAGAGCCTCTTTCGGCTACCTGGAGACACGGCGGTGGCG 289
QY 247 GGGCGGGCGGGGGCGGCTGATGACGACCGCGCGAGTGTGGCGTGAACATCTCC 306
Db 290 GGGCGGGCGGGGGCGGCTGATGACGACCGCGCGAGTGTGGCGTGAACATCTCC 349
QY 307 GCTTACCAAGAGACCCCGCGTACCTGGCCAGTGCCTGGCGCTCCGCGCCGCTGTG 366
Db 350 GCTTACCAAGAGACCCCGCGTACCTGGCCAGTGCCTGGCGCTCCGCGCCGCTGTG 409
QY 367 TACCCGCGCGCGGTGGCTGCTCTCATGTGTGGATGGCAACCGCGCCGAGACCTC 426
Db 410 TACCCGCGCGCGGTGGCTGCTCTCATGTGTGGATGGCAACCGCGCCGAGACCTC 469
QY 427 TACATGTTGACATGTCGCGAGTCTTCGCTGACGAGGACCCCGCCACGTACGTGG 486
Db 470 TACATGTTGACATGTCGCGAGTCTTCGCTGACGAGGACCCCGCCACGTACGTGG 529
QY 487 GACGGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGGTGGCGCCGAGCC 546
Db 530 GACGGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGGTGGCGCCGAGCC 589
QY 547 TATCGGAGGTGGAGCGAGATCTTGGCGGCTGGCAGTGGAGCGCTGTGAGGACT 606
Db 590 TATCGGAGGTGGAGCGAGATCTTGGCGGCTGGCAGTGGAGCGCTGTGAGGACT 649
QY 607 CGCAGTGGTGGTGGCGAGCGCTGGGGCGGCGAGCGGAGTCTATGACAGCC 666
Db 650 CGCAGTGGTGGTGGCGAGCGCTGGGGCGGCGAGCGGAGTCTATGACAGCC 709
QY 667 TTCAAGGGCGCTCGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACAAGTTG 726
Db 710 TTCAAGGGCGCTCGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACAAGTTG 769
QY 727 GACCCATGGCACTGCTGAGTCTGCGGGTACTGGACGAGGACCCCGGGTAGGGCT 786
Db 770 GACCCATGGCACTGCTGAGTCTGCGGGTACTGGACGAGGACCCCGGGTAGGGCT 829
QY 787 GTTGGTGGGAGCTGGGATCTTAACTCTGGACTCTGGTCTGAGTCTTCTAAGCAGC 846
Db 830 GTTGGTGGGAGTCTGGGATCTTAACTCTGGACTCTGGTCTGAGTCTTCTAAGCAGC 889
QY 847 CTGCGATACTGGGTAGCTTCAATGTGGAGCGGCTTGTGACAGTCTACTTCCACTGTGA 906
Db 890 CTGCGATACTGGGTAGCTTCAATGTGGAGCGGCTTGTGACAGTCTACTTCCACTGTGA 949
QY 907 TCCTGCATCAGCGGTCTCTAG----- 928
Db 950 TCCTGCATCAGCGGTCTCTAGCGCTATATAGGAATAACCTCTTGACGAGTTTCTTGAG 1009
QY 929 -----AATCTGCGAGGCGCCAGGAGCAGCGATGATGCCCTCAT 971
Db 1010 GCCTGTTACAACAGAAAGTTCTTGGGTACCCACTGTACTTTTGGGGATGACCGGACCTC 1069
QY 972 CTTGCGCC-----CGTGCAGGTACACCTCCAGGTCCCGCTGTAC 1012
Db 1070 ACCAACCGATGCTCAGCATGGGTATGTACAAAGTACACCTCCAGGTCCCGCTGTAC 1129
QY 1013 TCAGAGACGCGCTGCTCTTCTGCGGTGGCTGAGCCAGCAGACGCTGTGTCGAAGTCG 1072
Db 1130 TCAGAGACGCGCTGCTCTTCTGCGGTGGCTGAGCCAGCAGACGCTGTGTCGAAGTCG 1189
QY 1073 TACTTCCGTGA 1083
Db 1190 TACTTCCGTGA 1200
```

RESULT 7

ADL13693/c

ID ADL13693 standard; DNA; 231222 BP.

XX

```
AC ADL13693;
XX
DT 06-MAY-2004 (first entry)
XX
DE Osteoarthritis-associated polymorphic nucleotide #225.
XX
KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO2003054166-A2.
XX
PD 03-JUL-2003.
XX
PF 19-DEC-2002; 2002WO-US041225.
XX
PR 20-DEC-2001; 2001US-0342603P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Jones KA, Schafer A;
XX
DR WPI; 2003-559141/52.
XX
CC Determining susceptibility of an individual to joint space narrowing,
CC osteophyte development and/or joint pain comprises identifying whether
CC the individual has at least one polymorphism in a polynucleotide encoding
CC a protein.
PS Disclosure; SEQ ID NO 225; 297pp; English.
XX
CC The invention relates to a method of determining susceptibility of an
CC individual to joint space narrowing and/or osteophyte development and/or
CC joint pain comprising identifying whether the individual has at least one
CC polymorphism in a polynucleotide encoding at least one of the protein
CC listed in the specification. The methods, composition and agent are
CC useful for modulating the susceptibility of an individual to joint space
CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 231222 BP; 64006 A; 52087 C; 53478 G; 61650 T; 0 U; 1 Other;

Query Match 64.4%; Score 697.2; DB 10; Length 231222;
Best Local Similarity 99.6%; Pred. No. 5.6e-111;
Matches 699; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 CAGCAGAGCGGGCCCAAGCCCACTCTGCAGCCCGCCGCTGTCCGGCTTGCCTGCGGAGG 66
Db 149122 CAGCAGAGCGGGCCCAAGCCCACTCTGCAGCTGCGCTGTCCGGCTTGCCTGCGGAGG 149063
QY 67 GTGCTGACCATCGCTTCCGCTTGTCTATCTTGGGCTCATGACCTGGGCTTACCGCGCC 126
Db 149062 GTGCTGACCATCGCTTCCGCTTGTCTATCTTGGGCTCATGACCTGGGCTTACCGCGCC 149003
QY 127 GGGGTGCGCTGGCTCCGATCGCTACGGCTCTTGGCTTCCGGCTTACGGGCTTC 186
Db 149002 GGGGTGCGCTGGCTCCGATCGCTACGGCTTCTGGCTTCTACGGGCTTC 148943
QY 187 CTTTCAGCGCACCTGGTGGCGAGAGCCTCTTTCGGTACCTGGAGACCGCGGTGGCG 246
Db 148942 CTTTCAGCGCACCTGGTGGCGAGAGCCTCTTTCGGTACCTGGAGACCGCGGTGGCG 148883
QY 247 CGCGCGCGCGGGGCGGCTGATGACAGCCACCGCGCGAGTGTGGCGCTGACCATCTCC 306
XX
```

Db 14882 GCGCGCGCGCGCGCGCTGGATGACGACCGCGCGAGTGTGGCGCTGACCATCTCC 148823
 QY 307 GCCTACAGAGAGACCCCGCGTACCTGCGCCAGTGCCTTGGCTCGCCCGCGCCCTGCTG 366
 Db 148822 GCCTACAGAGAGACCCCGCGTACCTGCGCCAGTGCCTTGGCTCGCCCGCGCCCTGCTG 148763
 QY 367 TACCGCGCGCGCGCGCTGCGGCTCTCATGTGTGGATGCAACCGCGCGAGGACCTC 426
 Db 148762 TACCGCGCGCGCGCGCTGCGGCTCTCATGTGTGGATGCAACCGCGCGAGGACCTC 148703
 QY 427 TACATGTGTCATGTTCCCGAGGTCTTCGCTGACGAGGACCCCGCCACATGATGTGG 486
 Db 148702 TACATGTGTCATGTTCCCGAGGTCTTCGCTGACGAGGACCCCGCCACATGATGTGG 148643
 QY 487 GACGCACTACACAGCCCTTGGAAACCCGCGCGCGCGCGGTGGCGCCGAGCC 546
 Db 148642 GACGCACTACACAGCCCTTGGAAACCCGCGCGCGCGCGGTGGCGCCGAGCC 148583
 QY 547 TATCGGAGGTGGAGCGGAGATCCTGCGCGGCTGGCAGTGGAGGCGCTGTGAGGACT 606
 Db 148582 TATCGGAGGTGGAGCGGAGATCCTGCGCGGCTGGCAGTGGAGGCGCTGTGAGGACT 148523
 QY 607 CGCAGTGCCTGCTGCGCGCAGCGCTGGGCGCGGCGCGGCGGTGATGTACACAGCC 666
 Db 148522 CGCAGTGCCTGCTGCGCGCAGCGCTGGGCGCGGCGCGGCGGTGATGTACACAGCC 148463
 QY 667 TTCAGGCGCTCGGAGATTCGGTGAATACGTGCAAGTCTGT 708
 Db 148462 TTCAGGCGCTCGGAGATTCGGTGAATACGTGCAAGTGTAGT 148421

RESULT 8

ABZ76734
 ID ABZ76734 standard; cDNA; 1752 BP.
 XX AC ABZ76734;
 XX DT 01-APR-2000 (first entry)
 XX DE Mouse hyaluronan synthase 1 encoding cDNA SEQ ID NO:1.
 XX KW Mouse; hyaluronan synthase; HAS; HAS1; enzyme; ophthalmological;
 KW osteopathic; antiarthritic; gene therapy; angiogenesis inhibitor;
 KW hyaluronic acid; dihydrazide; bioconjugate; dry eye syndrome; eye;
 KW osteoarthritis; gene; ss.
 XX OS Mus sp.
 XX FH Key Location/Qualifiers
 XX CDS 1..1752
 XX FT /*tag= a
 XX FT /product= "hyaluronan synthase 1 (HAS1)"
 XX FN WO2003006068-A1.
 XX PD 23-JAN-2003.
 XX PF 10-JUL-2001; 2001WO-US021785.
 XX PR 10-JUL-2001; 2001WO-US021785.
 XX PA (CLEA-) CLEAR SOLUTIONS BIOTECH INC.
 XX PI Dehazya P, Chen W;
 XX DR WPI; 2003-221664/21.
 XX DR P-PSDB; ABP96028.
 XX PT Novel dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
 PT for treating dry eye syndrome, has derivatized hyaluronic acid
 PT crosslinked to nucleic acid encoding protein with hyaluronan synthase
 PT activity.
 XX

PS Claim 19; Page 56; 62pp; English.

XX The present invention describes a dihydrazide derivatised hyaluronic acid
 CC (HA)/nucleic acid (NA) bioconjugate (I), with derivatised HA linked to NA
 CC encoding protein with hyaluronan synthase (HAS) activity, where NA has
 CC sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659
 CC or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding
 CC proteins with 66-90% SI to a sequence of 593, 552 or 554 amino acids (see
 CC ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and
 CC antiarthritic activities, and can be used in gene therapy and as an
 CC inhibitor of angiogenesis, and as an inducer of expression of (HA) in
 CC human corneal epithelial cell. (I) is useful for transfecting a cell of
 CC an eye with nucleic acid, where the nucleic acid comprises (ABZ76734 to
 CC ABZ76736), by contacting the cell with (I) comprising the nucleic acid.
 CC (I) is useful for treating dry eye syndrome in an individual. (I) is
 CC useful in gene therapy applications for the treatment of a variety of
 CC medical conditions including dry eye syndrome or other medical conditions
 CC where an increase in the production of (HA) in the eye would be
 CC therapeutic (e.g., osteoarthritis of the articular joints). (I) is also
 CC useful for inhibiting angiogenesis for the treatment of macular
 CC degeneration or genes related to lipid biosynthesis that helps to restore
 CC the lipid component of the tear film, and as reagents for in vitro
 CC transformation of any cell, preferably a eukaryotic cell, more preferably
 CC a human eye cell. The present sequence encodes mouse HAS1 which is used
 CC in the exemplification of the present invention

XX SQ Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;

Query Match 62.6%; Score 677.8; DB 8; Length 1752;
 Best Local Similarity 84.5%; Pred. No. 1.4e-107;
 Matches 792; Conservative 0; Mismatches 127; Indels 18; Gaps 2;

QY 10 CAGGACGCGCGCAAGCCCACTCTCGAGCGCGCGCTGCTCCGGCCTGGCCGAGGGTG 69
 Db 7 CAGGACATGCGCAAGCCCTCAGAGCGCGCGTGTGCTCTGGCCTGGCCGAGCGCA 66
 QY 70 CTGACCATCGCTTCGCGCTGCTCATCTGGGCTCATGACCTGGGCTACGCGCGCGG 129
 Db 67 CTACGATCATCTTTGCGCTGCTCATCTGGGCTCATGACCTGGGCTACGCGCGAGGC 126
 QY 130 GTGCGCTGGCCTCCGATCGCTACGGCTCTCTGCGCTTCGCGCTTACGCGGCTTCTT 189
 Db 127 GTTCCTCTGGCTTCAGATCGTATGGACTCTGCGCTTTGGCCTCTATGGGCTTCTCT 186
 QY 190 TCAGCGACCTGTGGCGCAGAGCCTCTTGGCTACCTGGAGCACCGCGGCTGGCGCG 249
 Db 187 AGCGCACACCTAGTGGCAGAGCCTCTTGGCTTACCTGGAGCACCGAAGGCTGGCAGC 246
 QY 250 GCGCGCGG-----GGGCGCGTGGATGCGACCGACCGCGCGAGTGTGGCG 294
 Db 247 GCTGCGCGGCGCTCTTGGCGAAGGGGCCCCCTGGATCGGCCACTGACGCGAGGCTGGCA 306
 QY 295 CTGACCATCTCCGCTTACAGGAGGACCCCGCGTACCTGGCGCCAGTGCTGGCGTCCGCC 354
 Db 307 CTCACATCTCAGCCTACCAAGAGGATCCCGCTTACCTGGCGCAGTGCTTGACCTCCGCG 366
 QY 355 CGGCGCTCTGTGTACCGCGCGCGCTGGCGCTCTCATGTGTGGATGGCAACGCG 414
 Db 367 CGGCGCTTGTGTACCGCGCACAGAGGTTACGCGTGTCTCATGTGTGGAGCGCAACGCG 426
 QY 415 GCGGAGGACCTTACATGGTTCGACATGTTCCGGGAGGTCTTCGCTGACGAGGACCCCGCC 474
 Db 427 GCTGAGGATCTGTACATGGTGGACATGTTCCGAGAGTCTTCGCCGATGAGGACCCCGCC 486
 QY 475 ACGTACGTGTGGGACGGCAACTACCAACAGCCCTGGGAACCCCGCGCGG---CGGCGCG 531
 Db 487 ACTTATGTGTGGATGGCAACTACCAACAGCCCTGGGAACCCAGCGAGGCTTACGCGGCT 546
 QY 532 GTGGCGCGGAGCGCTTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAG 591
 Db 547 GTCGGTGAAGTGTCTTACCGGAGGTGGAGCGGAGGACCCCGGCGGCTTGGCGGTGGAG 606
 QY 592 GCGCTGTGAGGACTCGCAGGTGCTGTGTGCGCGCAGCGCTGGGGCGGCAAGCGCGAG 651

Db 607 GCGCTGGTGAACACGCGAGTGCCTGCTGGCTCAGCGTTGGGGCGCAACGCTGAG 666
Qy 652 GTCATGTATACACAGCTTCAAGCGCTTCGGAGATTGGTGGACTAGTGCAGGTCTGTGAC 711
Db 667 GTCATGTATACACAGCTTCAAGCGACTGGCGGACTCGTGGACTATGTCAGGTCTGTGAC 726
Qy 712 TCGGACACAAGGTTGGACCCCATGGCACTGTGGAGCTCGTGGGTACTGGACAGGAC 771
Db 727 TCAGACACAAGACTAGACCCCATGGCACTGTGGAGCTTGTGGAGTGTGGATGAAGAC 786
Qy 772 CCGCGGTAGGGCTGTGGTGGGACGTCGGATCTTAACTCTGGACTCTGGGTC 831
Db 787 CCGCGGTAGGGCTGTGGAGGGGATGTGAGGATCTTAACTCTGGACTCTGGGTC 846
Qy 832 AGCTTCTTAAGCAGCTGGTACTGGGTAGGCTTCAATGTGGAGGGCTTGTGAGAGC 891
Db 847 AGCTTCTTAAGCAGCTTTCGATACGGGTAGGCTTCAATGTGGAGGGCTTGTGAGAGC 906
Qy 892 TACTTCCACTGTGTATCTCTGCATCAGCGGTCTCTAG 928
Db 907 TACTTCCACTGTGTCTCTGCATCAGTGGTCTCTGG 943

RESULT 9

AAD59442
ID AAD59442 standard; DNA; 1752 BP.
AC AAD59442;
XX 18-DEC-2003 (first entry)
DT Mouse hyaluronan synthase (HAS) 1 DNA.
DE Mouse; hyaluronan synthase; HAS; dry eye syndrome;
KW eye cell; osteoarthritis; gene therapy; gene; db.
XX Mus sp.
OS
PH Key Location/Qualifiers
FT 1..1752
FT /*tag= a
FT /product= "Mouse hyaluronan synthase 1"
XX US2003087850-A1.
XX 08-MAY-2003.
XX 10-JUL-2001; 2001US-00902939.
XX 10-JUL-2001; 2001US-00902939.
XX (DEHA/) DEHAZYA P.
PA (CHEN/) CHEN W.
XX Dehazya P, Chen W;
PI WPI: 2003-755151/71.
DR P-PSDB; AAE39152.
XX Dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
PT comprising derivatized hyaluronic acid cross linked to nucleic acid
PT encoding hyaluronan synthase useful for treating dry eye syndrome.
XX Claim 1; Page 12-13; 31pp; English.
PS The invention relates to dihydrazide derivatized hyaluronic acid (HA)/
CC nucleic acid bioconjugate comprising derivatized HA cross linked to
CC nucleic acid encoding hyaluronan synthase (HAS). The invention is useful
CC for transfected an eye cell of an individual. It is useful for treating
CC dry eye syndrome and osteoarthritis of the particular joints. The
CC invention is also useful in gene therapy. The present sequence is mouse
CC HAS1 DNA

XX SQ Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;
Query Match 62.6%; Score 677.8; DB 10; Length 1752;
Best Local Similarity 84.5%; Pred. No. 1.4e-107; Indels 18; Gaps 2;
Matches 792; Conservative 0; Mismatches 127;
Qy 10 CAGGACGGCCCAAGCCCACTCTCTGACGCCCGCTGCTGCTGCGGCTGCGCCGAGGGTG 69
Db 7 CAGGACATGTCACCAAGCCCTCAGAGCAGCGGTTGCTGCTCTGCGCTGCGCCAGGACGA 66
Qy 70 CTGACCATCGCTTTCGCCCTGCTCATCTGCGGCTCATGACCTGCGGCTATCGCCCGCGG 129
Db 67 CTCAGATCATCTTTGGCTGCTCATCTGCGGCTCATGACCTGCGGCTCATCGCCAGGC 126
Qy 130 GTGCCGCTGGCTCCGATCGCTAGCGCTCTTGGCTTTCGGCTTTCAGCGGCTTCTCTT 189
Db 127 GTTCTCTGGCTTCAGATCGCTATGGAATCTCTGGCTTTCGGCTTTCAGCGGCTTCTCTC 186
Qy 190 TCAGCGCACCTGGTGGCGCAGAGCCTCTTTCGCTACCTGGAGCACCGCGGCTGCGCG 249
Db 187 AGCGCACACCTAGTGGCACAGAGCCTCTTTCGCTTACCTGGAGCACCGGAAGGCTGCG 246
Qy 250 GCGGCGCG-----GGGGCGCTGGATGCGACCCACCGCGCGCTGTGGCG 294
Db 247 GCTGCGCGCGCTCTTTCGCGAAGGGGCTTTCGCTTACCTGGAGCACCGGAAGGCTGCGCA 306
Qy 295 CTGACCATCTCGCTTACAGAGGACCCCGCTACCTGCGGCTGCTGCTGCGCTGCGCC 354
Db 307 CTCACCATCTCAGCTTACAGAGGATCCCGTTACCTGCGGCTGCTTTCAGCTTCTCGCG 366
Qy 355 CGCGCTCTGTATACCGCGCGCTGCGCTGCTCTCATGCTGGTGGATGCAACCGC 414
Db 367 CGCGCTTCTGTATACCGCACACGAGTTACCGCTGCTCATGCTGGTGGATGCAACCGC 426
Qy 415 GCGGAGCACTTACATGCTGCAATGTTTCGCGAGGCTTTCGCTGACGAGACCCCGCC 474
Db 427 GCTGAGGATCTGTATCATGCTGACATGTTCCGAGAAGTCTTCGCGGATGAGACCCCGCC 486
Qy 475 ACGTACGTGTGGGCGCAACTTACCACGACCCCTGGGAACCGCGCGCG--CGGCGCG 531
Db 487 ACTATGTGTGGGATGGCACTTACCATCAGCCCTGGGAACCGAGGCTTACCGGCGCT 546
Qy 532 GTGGCGCGCGGACCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAG 591
Db 547 GTCGCTGAAGTGTCTTACCGGAGGTGGAGCGGAGGACCCCGGCGGTTGCGCTGGAG 606
Qy 592 GCGCTGGTGAAGCTTCGCAAGTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCGAG 651
Db 607 GCGCTGGTGAAGACACGCGAGTGTGCTGGCTGCTCAGCGCTTGGGGCGGCAACGCTGAG 666
Qy 652 GTCATGTATACAGCTTCAAGCGCTCGGAGATTCGGTGGACTACGTCGAGCTGTGTGAC 711
Db 667 GTCATGTATACAGCTTTCAGGCACTGGGCGACTCCGTTGGAGCTACGTCGAGGCTGTGTGAC 726
Qy 712 TCGGACACAAGGTTGGACCCCATGGCACTGTGGAGCTCTGCGGCTTACGACGAGGAC 771
Db 727 TCAGACACAAGACTAGACCCCATGGCACTGTGGAGCTTGTGGAGTGTGTGATGAAGAC 786
Qy 772 CCGCGGTAGGGCTGTGGTGGGCACTGCGGATCTTAACTCTTGAACCTCTGAGACTCTGGGTC 831
Db 787 CCGCGGTAGGGCTGTGGAGGGGATGTGAGGATCTTAACTCTTGAACCTCTGAGACTCTGGGTC 846
Qy 832 AGCTTCTTAAGCAGCTGGTACTGCGGTAGGCTTCAATGTGGAGGGCTTGTGAGAGC 891
Db 847 AGCTTCTTAAGCAGCTTTCGATACCTGCGGTAGGCTTCAATGTGGAGGGCTTGTGAGAGC 906
Qy 892 TACTTCCACTGTGTATCTCTGCATCAGCGGTCTCTAG 928
Db 907 TACTTCCACTGTGTCTCTGCATCAGTGGTCTCTGG 943

AAT91655
ID AAT91655 standard; DNA; 2102 BP.
XX
AC AAT91655;
XX
DT 17-OCT-2003 (revised)
DT 19-DEC-1997 (first entry)
XX
DE Mouse hyaluronate synthase genomic DNA.
XX
KW Hyaluronate synthase; hyaluronic acid; polysaccharide; ds.
XX
OS Mus sp; (strain FM3A).
XX
FH Key Location/Qualifiers
FT CDS 49..1800
FT /tag= a
FT /product= "Hyaluronate_synthase"
XX
PN JP09224674-A.
XX
PD 02-SEP-1997.
XX
PF 26-FEB-1996; 96JP-00038336.
XX
PR 26-FEB-1996; 96JP-00038336.
XX
PA (KAGG) KAGAKU GIJUTSUUCHO CHOKAN KANBO.
XX
DR WPI; 1997-484102/45.
DR P-PSDB; AAW30704.
XX
XX Hyaluronate synthase isolated from mouse cells - useful for large-scale
PT production of hyaluronic acid.
XX
PS Claim 2; Page 10-13; 15pp; Japanese.
XX
CC A mouse-derived cell strain, FM3A P15-A, was cultured in Eagle minimal
CC medium containing 10 % heat-inactivated bovine serum, twice concentration
CC of amino acids and vitamins and penicillin and streptomycin at 37 degrees
CC Celsius. The culture was subjected to immobilised erythrocyte exclusion
CC to examine the extent of extracellular formation of hyaluronic acid
CC matrix. Cells which showed high formation were recovered and named FM3A
CC HAI. Cells synthesising no hyaluronic acid were recovered and named HAS-
CC The HAS- cell in which polyoma large r antigen was expressed was prepared
CC (HAS- P cell). Poly(A) + RNA was isolated from FM3A HAI and cDNA was
CC prepared and was used for constructing a library in HAS- P cells. Cells
CC having hyaluronic acid synthetic activity were selected from the
CC transformants and plasmid DNA was recovered and amplified in E.coli. The
CC resulting genomic DNA sequence codes for hyaluronate synthase having a
CC sequence of 583 amino acids which is used for large-scale production of
CC hyaluronic acid, a component of drugs and cosmetics. (Updated on 17-OCT-
CC 2003 to standardise OS field)
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 62.6%; Score 677.8; DB 2; Length 2102;
Best Local Similarity 84.5%; Pred. No. 1.4e-107;
Matches 792; Conservative 0; Mismatches 127; Indels 18; Gaps 2;
QY 10 CAGGACGGCGCCCAAGCCACCTCTGTCAGCGCGCGCTGCTCGGCTGGCGGAGGTG 69
DB 55 CAGGACATGCACCAAGCCCTCAGAGCAGCGCGGTGCTGCTGCTGGCTGGCGGAGCA 114
QY 70 CTGACCATCGCTTCGCGCTGCTCATCTCTGGGCTCATGACCTGGGCTAGCGCGCGGG 129
DB 115 CTCAGATCATCTTTGGCTGCTCATCTGGGCTCATGACCTGGGCTAGCGCGAGC 174
QY 130 GTGGCGCTGGCTCGCATCGCTACGGCTCTCTGGGCTTCGGCTTACGGGCTTCCTT 189
DB 175 GTTCTCTGGCTTCAGATCGCTATGCACTCTCGGCTTTGGCTCTATGGGCAATTCCTC 234
QY 190 TCAGGGCACCTGGTGGCGCAGAGCCTCTTCGGTACCTGGAGCACCGCGGGTGGCGGC 249

235 AGCGCACACCTAGTGGCACAGAGCCTCTTCCTTACCTGGAGCACCGAAGGTTGGCAGCG 294
250 CGCGCGCG-----GGGGCCGCTGATGACGACGCGCGCGCAGTGTGGCG 294
295 GCTGCGCGGCGCTCCTTTGGCGAAGGGGCGCTGATGGCGCACTGACGACGAGCGTGCA 354
295 CTGACCATCTCGGCTTACGAGAGGACCGCGGCTACCTGCGCCAGTGCCTGGCTCCGCC 354
355 CTCACCATCTCAGCCTTACCAAGAGGATCCCGCTTACCTGCGCCAGTGTTCGACCTCCCG 414
355 CGCGCTCTGTGTACCGCGCGCGCTGCGCGCTCTCTCATGTGTGGTGGTGAACACCGC 414
415 CGCGCTCTGTGTACCGCGCACACGAGGTTACGCGTCTCATGTGTGGTGGTGAACACCGC 474
415 CGCGAGGACCTCTACATGTGTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCC 474
475 GCTGAGGATCTGTACATGTTGACATGTTCCGAGAGTCTTTCGCGGATGAGGACCCCGCC 534
475 ACCTACGTGTGGAGGCGCACTACACAGCCCTGGGNAACCGCGCGCG-----CGGGCGCG 531
535 ACTTATGTGTGGGATGGCACTACCATCAGCCCTGGGAACACGAGCGAGGCTTACGGGCGCT 594
532 GTGGCGCGCGAGGCTTATCGGGAGGTGAGCGCGAGGATCTTCGCGCGCTGGCAGTGGAG 591
595 GTGCGTGAAGTGTCTACCGGAGGTGAGCGCGAGGACCCCGCGCGTGGCGTGGAG 654
592 GCCTGTGTGAGGACTCGCAGGTCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGGAG 651
655 GCGTGTGTGAGAACACGCGAGGTGCTGTGCGTGGCTCAGCGTTGGGGCGGCAACGCTGAG 714
652 GTCATGTACACAGCCTTCAGGCGCTCGGAGATTTCGGTGGACTACGTGCGAGTCTGTGAC 711
715 GTCATGTACACAGCTTTCAAGGGCACTGGGCGACTCCGTTGGACTACGTGCGAGTCTGTGAC 774
712 TCGGACACAAGGTTGGACCCCATCGCACTGCTGGAGCTCGTGGCGGTACTCGACGAGGAC 771
775 TCAGACACAGACTAGACCCCATGCGCACTGCTGGAGCTTGTGCGAGTGTGGATGAAGAC 834
772 CCGCGGTAGGGGCTGTTGGTGGGAGCTGCGGATCTCTTAAACCTCTGGACTCTCTGGGTC 831
835 CCGCGGTAGGGGCTGTTGGAGGGGATGTGAGGATCTTAAACCTCTGGACTCTCTGGGTC 894
832 AGCTTCTTAGCAGCCTTCGATATCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTACAGAC 891
895 AGCTTCTTAGCAGCTCTTCGATATCTGGGTAGCCTTCAATGTGGAGCAGGCTTGTACAGAC 954
892 TACTTCCACTGTGTATCTTCATCAGCGGCTCCTCTAG 928
955 TACTTCCACTGTGTGCTCTGATCAGTGGTCTCTCTGG 991

RESULT 11
AAZ10862
ID AAZ10862 standard; DNA; 2102 BP.
XX
AC AAZ10862;
XX
DT 19-OCT-1999 (first entry)
XX
DE Hyaluronate synthase coding sequence.
XX
KW Hyaluronate synthase; promoter; mouse; cell growth inhibitor; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 49..1800
FT /tag= a
XX
PN JP11196875-A.
XX
PD 27-JUL-1999.

QY 694 TA 695
 Db 661 TA 662

RESULT 15
 ID ABQ27385/c
 XX ABQ27385 standard; DNA; 662 BP.
 AC ABQ27385;
 XX 12-JUL-2002 (first entry)
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 13976.
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; da.
 XX Homo sapiens.
 OS
 XX WO200218632-A2.
 XX 07-MAR-2002.
 XX 01-SEP-2001; 2001WO-EP010074.
 XX 01-SEP-2000; 2000DE-01043826.
 PR 05-SEP-2000; 2000DE-01044543.
 XX (EPIC-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K, Guetig D;
 PI WPI; 2002-371829/40.
 XX
 XX Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridisation to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridised to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention
 XX
 SQ Sequence 662 BP; 255 A; 247 C; 80 G; 80 T; 0 U; 0 Other;

Search completed: March 13, 2005, 03:46:27
 Job time : 706.833 secs

Query Match 39.0%; Score 422; DB 6; Length 662;
 Best Local Similarity 77.3%; Pred. No. 1.3e-63;
 Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
 QY 34 GCAGCCCGCTGCTCCGGCTGCGCGAGGTGTCGACATCGCTTCGCTTCGCTC 93

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:08:52 ; Search time 4464.41 Seconds
(without alignments)
9233.835 Million cell updates/sec

Title: US-10-672-399-5
Perfect score: 1083
Sequence: 1 atgagacagcaggacgccc.....tccaagtcgtacttcgtga 1083

Scoring table: IDENTITY_NUC
Gapop 10_0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	922.4	85.2	2072	3	CR602106
2	920.8	85.0	1728	9	AY421569 Homo sapi
3	822.4	75.9	1058	4	BME43754 AGENCOURT
4	795.6	73.5	965	1	AL545447
5	679.4	62.7	1746	9	AY421571 Mus muscu
6	496	45.8	896	5	BX371636
7	479.4	44.3	3542	3	AK053726 Mus muscu
8	357.4	33.0	650	2	BB665418
9	311.4	28.8	1084	4	BME44718 AGENCOURT
10	297.4	27.5	551	4	BG707429
11	297.4	27.5	551	4	BG707608
12	268.2	24.8	508	9	CE377866
13	262.4	24.2	782	4	BI753116
14	259.4	24.0	610	5	BX327795
15	257.8	23.8	894	7	CF239758 AGENCOURT
16	253.6	23.4	668	8	BH057613
17	239	22.1	774	5	BP677403
18	236.4	21.8	4151	3	AK028582
19	221.8	20.5	895	5	BU911059 AGENCOURT
20	209	19.3	747	6	CD466132 Leukon2.2
21	206.8	19.1	604	4	BG016146
22	202.6	18.7	944	6	CA487249
23	193.8	17.9	495	7	CN664556 A0814F02-
24	189.6	17.5	739	9	AY421570 Pan trogl

25 187.8 17.3 918 7 CF220501
26 184.2 17.0 660 5 BP674137
27 179.8 16.6 739 7 CN304087
28 179.4 16.6 2891 3 AK079729
29 176 16.3 653 7 CV029395
30 174.8 16.1 857 5 BU906272
31 174.8 16.1 866 6 CA988122
32 172.4 15.9 848 6 CA987283
33 171 15.8 809 5 BP693553
34 170.6 15.8 748 5 BP699960
35 165.8 15.3 647 7 CN666988
36 165.4 15.3 781 5 BP693535
37 164.6 15.2 871 5 BU910414
38 163.6 15.1 920 7 CF220396
39 162.8 15.0 746 5 BP705367
40 162.8 15.0 840 5 BU911086
41 162.8 15.0 941 6 CA790329
42 162.6 15.0 853 6 CA988392
43 162.6 15.0 864 5 BU907701
44 162.6 15.0 898 5 BU906701
45 161 14.9 790 5 BP706190

ALIGNMENTS

RESULT 1
LOCUS CR602106 2072 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODI027YJ02 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR602106
VERSION CR602106.1 GI:50482913
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2072)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2072)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequenage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..2072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="CSODI027YJ02"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 85.2%; Score 922.4; DB 3; Length 2072;
Best Local Similarity 90.4%; Pred.No. 4.9e-174;
Matches 1038; Conservative 0; Mismatches 36; Indels 74; Gaps 2;
QY 10 CAGACGCGCCCAAGCCCACTCTCGACGCCCGCTCTCGGCTCGCCGAGGCTG 69
|||||
DB 58 CAGGACGCGCCCAAGCCCACTCTCGACGCCCGCTCTCGGCTCGCCGAGGCTG 117
|||||

Unpublished (1999)
Contact: Robert Strausberg, ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM12361 row: d column: 12
High quality sequence start: 29
High quality sequence stop: 615.
Location/Qualifiers
1. .1058
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5589083"
/lab_host="DH10B"
/clone_lib="NIH_MGC_125"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

JOURNAL
COMMENT

FEATURES
source

355	TACAGAGAGACCCCGGTACCTTGCGCACGTGCTGGCGTCCGCGCCGCTGCTCTAC	414
370	CCGCGGCGCGGTGGCGGCTCTCATGTGTGGATGGCAACGCGCGCAGGACCTCTAC	429
415	CCGCGGCGCGGTGGCGGCTCTCATGTGTGGATGGCAACGCGCGCAGGACCTCTAC	474
430	ATGGTGCACATGTTCCGCGAGGTCTTCGTGACGAGAACCCCGCACGTACGTGTGGAC	489
475	ATGGTGCACATGTTCCGCGAGGTCTTCGTGACGAGAACCCCGCACGTACGTGTGGAC	534
490	GGCAATTCACACAGCCCTGGGAACCCGCGGCGGGCGCGGTGGCGCGGAGCCTAT	549
535	GGCAATTCACACAGCCCTGGGAACCCGCGGCGGGCGCGGTGGCGCGGAGCCTAT	594
550	CGGAGGTGGAGCGCGAGGATCTTGGCGGCTGGCAGTGGAGGCGCTGGTGGAGACTCGC	609

```
Db 595 CCGGAGGTGGAGCGGAGGATCTCGGCGCTGGCAGTGGAGGCGCTGGTGGAGACTCGT 654
Qy 610 AGGTGCGTGTGCGTGGCAGCGCTGGGGCGCAAGCGGAGGTGATGTACACAGCCTTC 669
Db 655 ANGTGCGTGTGCGTGGCGCAGCGCTGGGGCGCAAGCGGAGGTGATGTACACAGCCTTC 714
Qy 670 AAGGCGCTCGGAGATTCGGTGACTACGTGACAGTCTGTGACTCGGACACAGGTTGGAC 729
Db 715 AAGGCGCTCGGAGATTCGGTGACTACGTGACAGTCTGTGACTCGGACACAGGTTGGAC 774
Qy 730 CCCATGGCACTGCTCGAGCTCGTGGGGTACTGGACGAGGACCCCGGGTAGGGCTGTT 789
Db 775 CCCATGGCACTGCTCGAGCTCGTGGGGTACTGGACGAGGACCCCGGGTAGGGCTGTT 834
Qy 790 GGTGGGACGTGGGATCTTAACCTCTGGACTCTGGTCTGAGTCTGCTTAAAGACGCTG 849
Db 835 GNGGGGACGTGGGAACTTAACCTCTGGACTCTGGTCTGAGTCTGCTTAAAGCAACCTG 894
Qy 850 CGAT-ACTGGGTAGCTTCAAT-GTGGAGCGGGCTTG 884
Db 895 CGAATACTGGGTAGCTTCAATGGTGGAAACGGGCTTG 931
```

RESULT 4
AL545447 965 bp mRNA linear EST 25-MAR-2004
LOCUS AL545447 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1027YJ02 5-PRIME, mRNA sequence.

ACCESSION AL545447

VERSION AL545447.3 GI:45745929

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Li W.B., Gruber, C., Jessee, J. and Pollayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On Feb 15, 2001 this sequence version replaced gi:31267282.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9502.r

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?&CS0D1027DE01QPI&c=9502.r>.

FEATURES

Location/Qualifiers

1..965
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1027YJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 73.5%; Score 795.6; DB 1; Length 965;

Best Local Similarity 97.4%; Pred. No. Ie-148;

Matches 851; Conservative 9; Mismatches 8; Indels 6; Gaps 5;

Qy 10 CAGGACGGCCCAAGCCCACTCTCGACCGCGCGCTGCTCGGCTGGCCGGAGGGTG 69

```
Db 58 CAGGACGGCCCAAGCCCACTCTCGACCGCGCGCTGCTCGGCGCTGGCCCGGA -GGTG 116
Qy 70 CTGACCATCGCTTTCGCCCTGCTCATCTGGGCTCATGACTGGGCTTACGCCCGGG 129
Db 117 CTGACCATCGCTTTCGCCCTGCTCATCTGGGCTCATGACTGGGCTTACGCCCGGG 176
Qy 130 GTGCCGTGGGCTCCGATCGCTAGCGGCTCTTGGCTTTCGGCTTACCGGGCTTCTCTT 189
Db 177 GTGCCGTGGGCTCCGATCGCTAGCGGCTCTTGGCTTTCGGCTTACCGGGCTTCTCTT 236
Qy 190 TCAGCGCACTGTGTGGCGACAGGCTCTTCCGCTACTTGGAGCAACCGCGGTTGGCGCG 249
Db 237 TCAGCGCACTGTGTGGCGACAGGCTCTTCCGCTACTTGGAGCAACCGCGGTTGGCGCG 296
Qy 250 CGCGCGGGGGCGCTGGATGAGCCACCGCGGCGAGTGTGGCTGACCATCTCCGCC 309
Db 297 CGCGCGGGGGCGCTGGATGAGCCACCGCGGCGAGTGTGGCTGACCATCTCCGCC 356
Qy 310 TACCAGGAGGACCCCGGCTACTCGCGCAGTGCCTGGGCTCCGCCGCGCTGCTGTAC 369
Db 357 TACCAGGAGGACCCCGGCTACTCGCGCAGTGCCTGGGCTCCGCCGCGCTGCTGTAC 416
Qy 370 CGCGCGCGGCTGCGGCTCTCATGTGTGGATGGCAACCGCGCGGAGGACCTCTTAC 429
Db 417 CGCGCGCGGCTGCGGCTCTCATGTGTGGATGGCAACCGCGCGGAGGACCTCTTAC 476
Qy 430 ATGTGTGACATGTTCCGGGAGGTCTTGTGACAGAGACCCCGGCACTGCTGTGGGAC 489
Db 477 ATGTGTGACATGTTCCGGGAGGTCTTGTGACAGAGACCCCGGCACTGCTGTGGGAC 536
Qy 490 GGCAACTACCAACAGCCCTGGGAAACCGCGCGGCGGCTGGGCGCGGAGCCTTAT 549
Db 537 GGCAACTACCAACAGCCCTGGGAAACCGCGCGGCGGCTGGGCGCGGAGCCTTAT 596
Qy 550 CGGAGGTGGAGGCGGAGGATCTTGGGCGGCTGGCAGTGGAGGCGCTGGTGAAGACTCG 609
Db 597 CGGAGGTGGAGGCGGAGGATCTTGGGCGGCTGGCAGTGGAGGCGCTGGTGAAGACTCG 656
Qy 610 AGTGTGCTGTGCTGGGCGGAGCGCTGGGCGGCGGAGCGGAGGTCATGTACACAGCCTTC 669
Db 657 AGTGTGCTGTGCTGGGCGGAGCGCTGGGCGGCGGAGCGGAGGTCATGTACACAGCCTTC 716
Qy 670 AAGCGCTCGGAGATTCGGTGGACTACTGTGAGGTCTGTGACTCGGACACAAAGGTTGGAC 729
Db 717 AAGCGCTCGGAGATTCGGTGGACTACTGTGAGGTCTGTGACTCGGACACAAAGGTTGGAC 776
Qy 730 CCCATGGCACTGTGGAGCTGTCGGGTACTGGACAGGACCCCGGGTAGGGCTGTT 789
Db 777 CCCATGGCACTGTGGAGCTGTCGGGTACTGGACAGGACCCCGGGTAGGGCTGTT 836
Qy 790 GGTGGGACGTGGGATCCTTAACCTTGGACTCTGGGTCAGCTTCTTAAGCAGCCTG 849
Db 837 GGTGGGACGTGGGATCCTTAACCTTGGACTCTGGGTCAGCTTCTTAAGCAGCCTG 892
Qy 850 CGTACTGGGTAGCTTCAATGTGGAGCGGCTT 883
Db 893 CGTACTGGGTAG-CTTAAATGTGGAGCGGCTT 925
```

RESULT 5

LOCUS AY421571

DEFINITION Mus musculus Hs1 gene, VIRTUAL TRANSCRIPT, partial sequence,

genomic survey sequence.

ACCESSION AY421571

VERSION AY421571.1 GI:39748430

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1746)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,

AY421571 1746 bp DNA linear GSS 17-DEC-2003
Mus musculus Hs1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

896	GCACGTGGCTGGGGTCCGCCCGCCCTTGCTGTACCGGGGGCCGGTGGGGCGCTCTCTCA	831
395	TGTTGGTGGATGGCAACCGCGCCGAGGACCTTACATCGTTCGACATGTTCCGCGAGGTCT	454
836	TGGT-GTGGAGGACACCGCGCCGAGGACTTTACAATGGTCGACA-GTTCCCGCAGGTCT	779
455	TCGCTGACGAGGACCCCGCCACAGTACGTGTGGGACGGCAACTACACAGCCCTTGGG-AA	513
778	TCGCTGACGAGGACCCCGCCACAGTACGTGT-GGACGGCAACTACACAGCCCTTGGGAAA	720
514	CCCGCGCGCGCGCGGTGGCGCGCGAGCCCTATCGGAGGTGGAGCGGAGGATCCT	573
719	CCNCNCGNCGGCGCNCGGTGGGCGCCGAGCCCTATCGGAGGTGGAGCGGATCCT	660
574	GGCGCGCTGGCAGTGGAGGGCGCTGGTGGAGACTCGCAGGTGCGTGTGCTGGCGCAGCGC	633
659	GGNCGGTGGCAGTGGAGGGCGCTGGTGGAGACTCGCAGGTGCGTGTGCTGGCGCAGCGC	600
634	TGGGGCGCAAGCGCGAGGTCAATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGGAC	699
599	TGGGGCGCAAGCGCGAGGTCAATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGGAC	540
694	TAGTGCAGCTCTGTGACTCGGACACAAAGTTGGAACCCCATGGGCACCTGCTGGAGCTCGT	753
539	TACGTGCAGCTCTGTGACTCGGACACAAAGTTGGAACCCCATGGGCACCTGCTGGAGCTCGT	480
754	CGGTACTCGACGAGGACCCCGGGTAGGGCTGTGTGTGGGACGTGCGGATCCTTAAAC	813
479	CGGTACTCGACGAGGACCCCGGGTAGGGCTGTGTGTGGGACGTGCGGATCCTTAAAC	420
814	CCTCTGACCTCTGGGTGAGCTTCTTAAGCAGCCTGCGATATCGGTAGCGCTTCAATGTG	873
419	CCTCTGACCTCTGGGTGAGCTTCTTAAGCAGCCTGCGATATCGGTAGCGCTTCAATGTG	360
874	GAGGGGCTGTGACAGACTACTTCCACTGTGTATCCTGCATCAGCGGTCTCTTAG-----	928
359	GAGGGGCTGTGTGACAGACTACTTCCACTGTGTATCCTGCATCAGCGGTCTCTTAGGCCTA	300
929	-----AATCCTGCGCC	938
299	TATAGGATTAACCTTTGACAGCTTCTTGAGGCGCTGTACACAGAGTTCTCGGT	240
939	AGGCCCCAGGAGACGCGATGATGCCCCCTCATTCCTCGCCCC-----	980
239	ACCCACTGTACTTTTGGGGATGACCGGCACCTCTACCAACCGCATGTCTCAGCATGGTTAT	180
981	-CGTGCAGGTACACCTTCAGGTCCCGTGTCTACTCAGAGAGCGCCCTGCTCTTCTCGGG	1039
179	GCTACCAAGTACACCTTCAGGTCCCGTGTCTACTCAGAGAGCGCCCTGCTCTTCTCGGG	120
1040	TGGCTGAGCCAGCAGACACGCTGGTCCAAGTCTGACTTTCGGTGA	1083
119	TGGCTGAGCCAGCAGACACGCTGGTCCAAGTCTGACTTTCGGTGA	76
RESULT 7		
AK053726		
LOCUS	3542 bp	mRNA
DEFINITION	Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30302P16	product:hyaluronan synthase1, full insert sequence.
ACCESSION	AK053726	
VERSION	AK053726.1	GI:26343670
KEYWORDS	HTC; CAP trapper.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1	
TITLE	Carninci, P. and Hayashizaki, Y.	
JOURNAL	High-efficiency full-length cDNA cloning	
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)	
	99279253	

RESULT 8	BB665418	LOCUS	EST 26-OCT-2001	linear	650 bp	mRNA
<p>primed with a primer 15'- GAGAGAGAGCGGCCCAACTCGATGTTTTTTTTTTN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length b transcriptase and subsequently enriched for full-length b</p>						

366 CCACACTGTACTTTTGGGGAATGACCGGCACTCAACCAACGCATGCTCAGCATGGGTATG 42
 981 CGTGCAGGTACACCTCCAGGTCCCGTGTCTACTCAGAGAGCCCTCGTCTCCTCGGT 104
 426 CTACCAAGTACACCTCCAGGTCCCGTGTCTACTCAGAGAGCCCTCGTCTCCTCGGT 485
 1041 GGCTGAGCCAGCAGACACGCTGGTCCAAAGTCGTACTTCCGTGA 1083
 486 GGCTGAGCCAGCAGACACGCTGGTCCAAAGTCGTACTTCCGTGA 528
 RESULT 11
 BG707608
 LOCUS 602670672F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793704 5', mRNA linear EST 07-MAY-2000
 DEFINITION mRNA sequence.
 ACCESSION BG707608
 VERSION BG707608.1 GI:13984125
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 551)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Straubeberg, Ph.D.
 Email: csapbs-x@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: L1AM10678 row: p column: 12
 High quality sequence stop: 551.
 Location/Qualifiers
 1..551
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4795643"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_96"
 /notes="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
 FEATURES
 source
 1..551
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4793704"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_96"
 /notes="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 27.5%; Score 297.4; DB 4; Length 551;
 Best Local Similarity 79.0%; Pred. No. 3.3e-49;
 Matches 413; Conservative 0; Mismatches 36; Indels 74; Gaps 2;
 635 GGGGCGGCAAGCCGAGGTCAATGACAGCCCTTCAAGCGCTCGAGATTTCGGTGACT 694
 6 GGGGCGGCAAGCCGAGGTCAATGACAGCCCTTCAAGCGCTCGAGATTTCGGTGACT 65
 695 ACCTGACGCTCTGACTCGGACACAGGTTGGACCCCATGCGACTCTGGAGCTCGTGC 754
 66 ACCTGACGCTCTGACTCGGACACAGGTTGGACCCCATGCGACTCTGGAGCTCGTGC 125
 755 GGCTACTGACGAGGACCCCGGGTAGGGGCTGTTGGTGGGACGTCGGATCCTTAACC 814
 126 GGCTACTGACGAGGACCCCGGGTAGGGGCTGTTGGTGGGACGTCGGATCCTTAACC 185
 815 CTCTGACCTCTGGGTCAAGCTTCTTAAGCAGCGCTCGATATCTGGGTAGCCTTCAATGG 874
 186 CTCTGACCTCTGGGTCAAGCTTCTTAAGCAGCGCTCGATATCTGGGTAGCCTTCAATGG 245
 875 AGCGGCTTGTACAGCTACTTCCATCTGTATCTTCGATCAGCGTCTCTAG----- 928
 246 AGCGGCTTGTACAGCTACTTCCATCTGTATCTTCGATCAGCGTCTCTAGAGCCCTAT 305
 929 -----AATCCTCGCCA 939
 306 ATAGGAATAACCTCTTTCAGCAGTCTTCTTGGGCTGGTACAAACGAGTTCCTGGGTA 365
 940 GGCCCCGAGGAGCACGCGATGATGCCCTCATTTCTCGCCCC----- 980

Db 126 GGGTACTGGACGAGACCCCGGGTAGGGCTGTTGGTGGGACGTGCGGATCCTTAACC 185
QY 815 CTCTGGACTCTGGGTGAGCTTCCTAAGCAGCTCGGATCTGGGTGAGCTTCATATGG 874
Db 186 CTCTGGACTCTGGGTGAGCTTCCTAAGCAGCTCGGATCTGGGTGAGCTTCATATGG 245
QY 875 AGCGGGCTTGTGACAGCTACTTCCACTGTGTATCTGTCATCAGCGGTCTCTAG----- 928
Db 246 AGCGGGCTTGTGACAGCTACTTCCACTGTGTATCTGTCATCAGCGGTCTCTAGCGCTAT 305
QY 929 -----AATCTCTGCCCA 939
Db 306 ATAGGAATACTCTTTCAGCAGTTCCTTTCAGGCTGGTACAAACAGAGTTCCTGGTA 365
QY 940 GGCCCCAGGAGCAGCAGATGATGATCCCTCATCTTCGCCCC----- 980
Db 366 CCCACTACTTTTGGGGATGACCGGCACCTCACCAACCGCATGCTCAGCATGGGTTATG 425
QY 981 CGTCAGGTACACTCCAGGTCCGCTGCTACTCAGACGCGCTCGTCTCTCTCGGT 1040
Db 426 CTACCAAGTACACTCCAGGTCCGCTGCTACTCAGACGCGCTCGTCTCTCTCGGT 485
QY 1041 GCGTGAGCCAGCAGACGCTGGTCCAAAGTCTGCTACTTCCGTGA 1083
Db 486 GCGTGAGCCAGCAGACGCTGGTCCAAAGTCTGCTACTTCCGTGA 528

RESULT 12
CE377866
LOCUS
DEFINITION
tigr-gss-dog-1700036208906 Dog Library Canis familiaris genomic,
genomic survey sequence.
ACCESSION
CE377866
VERSION
CE377866.1 GI:36605926
KEYWORDS
GSS.
SOURCE
Canis familiaris (dog)
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1 (bases 1 to 508)
Kirkness,B.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
COMMENT
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
1. .508
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

FEATURES
source
1. .508
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 24.8%; Score 268.2; DB 9; Length 508;
Best Local Similarity 91.1%; Pred. No. 2.2e-43;
Matches 285; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 397 GTGTGGATGGCAACCGCGGAGACCTCTACATGGTCGACATGTTCCGCGAGGTCTTC 456
|||||

Db 1 GTGTGGACGCGCAACCGCCCGAGGACCTCTACATGTTGGACATGTTCCGCGAGGTCTTC 60
QY 457 GCTGACGAGACCCCGCCACGTCAGTGTGGACGGCAACTACACGAGCCCTGGGAACCC 516
Db 61 GCCAGCAGGACCCCGGCACTACGTGTGGACGGCAACTACACGAGCCCTGGGAGCCG 120
QY 517 GCGCGCGCGCGCGGTGGCGCGGAGCCCTATCCGAGGTGGAGCGGAGGATCCTGGG 576
Db 121 GCGCGCGCGCGCGGAGCGCGGCGGCGCTTACCGGAGGTGGAGCGCGAAGACCCCGG 180
QY 577 CGGTGCGAGTGGAGGCGCTGTTGAGGACTCGCAGGTGCGGTGCGCGCAGCGCTGG 636
Db 181 CGGTGCGGTGGAGCGCTGTTGAGGACGACAGGTGCGGTGCGGTGGCGCAGCGCTGG 240
QY 637 GCGGCAAGCGCGAGGTTCATGTACAGCCCTTCAAGCGCTCGAGATTCGGTGGACTAC 696
Db 241 GCGGCAAGCGCGAGGTTCATGTACAGCGCTTCAAGCGCTCGATGACTCTGTGGACTAC 300
QY 697 GTGCAGGTCTGTG 709
Db 301 GTGCAGGTGAGTG 313

RESULT 13
BI753116
LOCUS
DEFINITION
603025961F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196399 5',
mRNA sequence.
ACCESSION
BI753116
VERSION
BI753116.1 GI:15744694
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 782)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: rgs@nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11491 row: j column: 16
High quality sequence stop: 780.
Location/Qualifiers
1. .782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5196399"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

FEATURES
source
1. .782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5196399"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match 24.2%; Score 262.4; DB 4; Length 782;
Best Local Similarity 77.6%; Pred. No. 3.2e-42;
Matches 377; Conservative 0; Mismatches 36; Indels 73; Gaps 2;

```
QY 671 AGGCGCTCGAGATTCCGTGGAGTACGTGCGAGGTCGTGACTCGGACACAAAGTTGGACC 730
Db 6 AAGCGCTCGAGATTCCGTGGAGTACGTGCGAGGTCGTGACTCGGACACAAAGTTGGACC 65
QY 731 CCATCGGCACTGCTGAGCTCGTGGGCTACTGAGAGAGAGACCCCGGGTAGGGGCTGTG 790
Db 66 CCATCGGCACTGCTGAGCTCGTGGGCTACTGAGAGAGAGACCCCGGGTAGGGGCTGTG 125
QY 791 GTGGGAGCTGCGGATCTCTTAACCCCTCTGGACTCTCTGGGTCAGCTTCCTAAGCAGCCTGC 850
Db 126 GTGGGAGCTGCGGATCTCTTAACCCCTCTGGACTCTCTGGGTCAGCTTCCTAAGCAGCCTGC 185
QY 851 GATACCTGGTAGCCTTCAATGTGGAGCGGGCTTGTTCAGAGCTACTTCCACTGTGTATCCT 910
Db 186 GATACCTGGTAGCCTTCAATGTGGAGCGGGCTTGTTCAGAGCTACTTCCACTGTGTATCCT 245
QY 911 GCATCAGCGGTCCTCTA----- 927
Db 246 GCATCAGCGGTCCTCTTAGGCCATATAGGAATAAAGCTCTTGAGAGGCTTCTTGAGGCT 305
QY 928 GAATCCTGCCAGGCCCGGAGGACGCGATGATGCCCTCATTCCTCGCCCC----- 980
Db 306 GGTACACACAGAGTCTCTGGGTACCCACTGTACTTTGGGGATGACCGGACCTCACCA 365
QY 981 -----CCTGAGGTACACCTCCAGTCCCGTCCGCTGCTACTCAGA 1017
Db 366 CCGCATGCTCAGCATGGGTTATGTACCAAGTACACCTCCAGGTCCTGCTACTCAGA 425
QY 1018 GAGCGCTCGCTCTTCCGTGGTGTGCTGAGCCAGAGACAGCTCGTCCAGTCTACTTT 1077
Db 426 GAGCGCTCGCTCTTCCGTGGTGTGCTGAGCCAGAGACAGCTCGTCCAGTCTACTTT 485
QY 1078 CCGTGA 1083
Db 486 CCGTGA 491

RESULT 14
LOCUS BX327795
DEFINITION BX327795 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS01027YJ02 5-PRIME, mRNA sequence.
ACCESSION BX327795
VERSION BX327795.2 GI:46277978
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 610)
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT On May 2, 2003 this sequence version replaced gi:30340686.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqefgen@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS0BAG028ZH06_CS02640_1&c=9502.r

FEATURES
source
1. .610
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01027YJ02"

/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 24.0%; Score 259.4; DB 5; Length 610;
Best Local Similarity 97.4%; Pred. No. 1.3e-41;
Matches 263; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 659 ACACAGCCTTCAAGCGCTCGAGATTCCGTGGAGTACGTGAGGTCGTGACTCGGACA 718
Db 1 ACATGCTCTTAAGNCGCTCGAGATTCCGTGGAGTACGTGAGGTCGTGACTCGGACA 60
QY 719 CAAGTTGGAGCCCATGACCTGCTGGAGCTCGTGGGCTACTGGAGGACACCCCGGG 778
Db 61 CAAGTTGGAGCCCATGACCTGCTGGAGCTCGTGGGCTACTGGAGGACACCCCGGG 120
QY 779 TAGGGGCTGTTGGTGGGAGCTGCGGATCCCTTAACCTCTCGACTCTCTGGGTCAGCTTCC 838
Db 121 TAGGGGCTGTTGGTGGGAGCTGCGGATCCCTTAACCTCTCGACTCTCTGGGTCAGCTTCC 180
QY 839 TAAGCAGCTCGGATACCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCC 898
Db 181 TAAGCAGCTCGGATACCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCC 240
QY 899 ACTGTGATCTCTGATCAGCGGCTCTCTAG 928
Db 241 ACTGTGATCTCTGATCAGCGGCTCTCTAG 270

RESULT 15
LOCUS CF239758
DEFINITION CF239758 894 bp mRNA linear EST 05-AUG-2003
AGENCOURT 15100464 NICHD_XGC_Emb6 Xenopus tropicalis cDNA clone
IMAGE:6994934 5', mRNA sequence.
ACCESSION CF239758
VERSION CF239758.1 GI:33442966
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
REFERENCE 1 (bases 1 to 894)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14677 Row: m column: 13
High quality sequence stop: 699.
Location/Qualifiers
1. .894
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6994934"
/tissue_type="neurula"
/dev_stage="embryo, stages 14-19"
/lab_host="DH10B (phage-resistant)"
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:01:16 ; Search time 213.323 Seconds
(without alignments)
8307.054 Million cell updates/sec

Title: US-10-672-399-5
Perfect score: 1083
Sequence: 1 atgagacagcagcagccgcccc.....tccaaagtcgtacttcgtga 1083

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A-COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B-COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A-COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B-COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUS-COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	933	86.1	2116	3	US-08-635-552A-1 Sequence 1, Appli
2	929.8	85.9	2087	4	US-09-949-016-5659 Sequence 5659, Ap
3	920.6	85.0	2117	3	US-09-155-768-1 Sequence 1, Appli
4	697.2	64.4	14862	4	US-09-949-016-17401 Sequence 17401, A
5	396.4	36.6	601	4	US-09-949-016-201775 Sequence 201775,
6	273	25.2	1662	4	US-09-949-016-3728 Sequence 3728, Ap
7	236.4	21.8	1665	4	US-08-812-008-31 Sequence 31, Appli
8	229.4	21.2	601	4	US-09-949-016-201780 Sequence 201780,
9	179.4	16.6	2947	4	US-08-675-499A-1 Sequence 1, Appli
10	179.4	16.6	2947	4	US-08-812-008-1 Sequence 1, Appli
11	167.6	15.5	9871	4	US-09-949-016-15470 Sequence 15470, A
12	153.6	14.2	3003	4	US-09-949-016-737 Sequence 737, App
13	137.4	12.7	601	4	US-09-949-016-132199 Sequence 132199,
14	82.8	7.6	6645	4	US-09-902-540-8819 Sequence 8819, Ap
15	82.8	7.6	10210	4	US-09-902-540-938 Sequence 938, App
16	69.2	6.4	1320	4	US-09-724-797-85 Sequence 85, Appl
17	66.6	6.1	3957	4	US-10-237-551-193 Sequence 193, App
18	66.6	6.1	154746	4	US-09-827-688-8 Sequence 8, Appli
19	66.6	6.1	154746	4	US-09-827-688-8 Sequence 8, Appli
20	66.2	6.1	1614	4	US-09-616-289-45 Sequence 45, Appl
21	66.2	6.1	12425	3	US-09-616-289-50 Sequence 50, Appl
22	62	5.7	1292	3	US-08-483-533-37 Sequence 37, Appl
23	62	5.7	1292	3	US-09-283-471A-37 Sequence 37, Appl
24	60.6	5.6	4257	2	US-08-690-473-1 Sequence 1, Appli
25	60.6	5.6	4257	3	US-09-259-821A-1 Sequence 1, Appli
26	60.6	5.6	4257	3	US-08-843-659-1 Sequence 1, Appli
27	60.6	5.6	4257	4	US-09-825-288A-1 Sequence 1, Appli

c	28	60.6	5.6	12001	1	US-08-458-568A-11	Sequence 11, Appl
	29	60.4	5.6	32176	4	US-09-949-016-12479	Sequence 12479, A
	30	60.2	5.6	2900	1	US-08-034-650-9	Sequence 9, Appli
	31	60.2	5.6	2900	1	US-08-449-015-9	Sequence 9, Appli
	32	59.6	5.5	1647	4	US-09-902-540-4211	Sequence 4211, Ap
	33	59.6	5.5	21330	4	US-09-902-540-1209	Sequence 1209, Ap
c	34	59.4	5.5	1926	3	US-09-249-585A-4	Sequence 4, Appli
c	35	59.4	5.5	1931	2	US-09-130-114-2	Sequence 2, Appli
	36	59.4	5.5	441529	3	US-09-103-840A-1	Sequence 1, Appli
	37	59.2	5.5	1051	2	US-08-865-273-1	Sequence 1, Appli
	38	59.2	5.5	1051	3	US-09-385-174-1	Sequence 31, Appl
	39	59.2	5.5	1182	4	US-09-673-198-31	Sequence 31, Appl
	40	59.2	5.5	1361	4	US-09-614-912-37	Sequence 37, Appl
	41	59	5.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	42	58.4	5.4	5808	4	US-09-902-540-794	Sequence 794, App
	43	58.2	5.4	967	4	US-09-902-540-7886	Sequence 7886, Ap
	44	57.8	5.3	8604	4	US-09-902-540-5331	Sequence 5331, Ap
	45	57.8	5.3	33399	4	US-09-902-540-1260	Sequence 1260, Ap

ALIGNMENTS

RESULT 1
US-08-635-552A-1
; Sequence 1, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Brieskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
US-08-635-552A-1

Query Match 86.1%; Score 933; DB 3; Length 2116;
Best Local Similarity 90.6%; Pred. No. 9.6e-176;
Matches 1048; Conservative 0; Mismatches 35; Indels 74; Gaps 2;
QY 1 ATGAGACAGCAGCAGCGCCCAAGCCGCTCTCTGAGCCGCGCTGCTCCGGCTGGCC 60
|||||

Db 36 ATGAGACAGCAGCGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGCGCTGGCC 95
Qy 61 CGGAGGCTGTGACCATGCGCTTCGCGCTGCTCATCTGCGGCTCATGACCTGGCCCTAC 120
Db 96 CGGAGGCTGTGACCATGCGCTTCGCGCTGCTCATCTGCGGCTCATGACCTGGCCCTAC 155
Qy 121 GCGCGCGGCTGCGCTGCGCTCCGATCGTACGCGCTCTGCGCTTCGCGCTCTACGCG 180
Db 156 GCGCGCGGCTGCGCTGCGCTCCGATCGTACGCGCTCTGCGCTTCGCGCTCTACGCG 215
Qy 181 GCGCTTCCTTTCAGCGCACCTGCTGCGCAGAGCCCTTCCTGCGTACCTGGAGCACCGCGG 240
Db 216 GCGCTTCCTTTCAGCGCACCTGCTGCGCAGAGCCCTTCCTGCGTACCTGGAGCACCGCGG 275
Qy 241 GTGGCGGCGCGCGCGCGCGCGCTGATGAGCGCACCGCGCGAGTGTGGCGCTGAC 300
Db 276 GTGGCGGCGCGCGCGCGCGCGCTGATGAGCGCACCGCGCGAGTGTGGCGCTGAC 335
Qy 301 ATCTCGGCTTACAGAGAGACCCCGGTACCTGCGCGAGTGTGGCTGCGCGCGCGCC 360
Db 336 ATCTCGGCTTACAGAGAGACCCCGGTACCTGCGCGAGTGTGGCTGCGCGCGCGCC 395
Qy 361 CTGCTGTATACCGCGCGCGCGCTGCGCGCTCTCATGCTGTGGATGGCAACCGCGCGAG 420
Db 396 CTGCTGTATACCGCGCGCGCGCTGCGCGCTCTCATGCTGTGGATGGCAACCGCGCGAG 455
Qy 421 GACCTCTACATGCTCGAATGCTTCGCGAGGCTTCCTGCTGAGAGAGACCCCGCGCACGTAC 480
Db 456 GACCTCTACATGCTCGAATGCTTCGCGAGGCTTCCTGCTGAGAGAGACCCCGCGCACGTAC 515
Qy 481 GTGTGGAGCGCACTACCAAGCCCTGGGAACCCCGCGCGCGCGCGGTGGCGCC 540
Db 516 GTGTGGAGCGCACTACCAAGCCCTGGGAACCCCGCGCGCGCGCGGTGGCGCC 575
Qy 541 GGAGCGCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGGCACTGGAGCGCGCTGTG 600
Db 576 GGAGCGCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGGCACTGGAGCGCGCTGTG 635
Qy 601 AGGACTCGAGGT 660
Db 636 AGGACTCGAGGT 695
Qy 661 ACAGCTTTCAGGCGCTCGAGATTCGCTGAGTACGCTGAGTCTGAGTCTCGAGACA 720
Db 696 ACAGCTTTCAGGCGCTCGAGATTCGCTGAGTACGCTGAGTCTGAGTCTCGAGACA 755
Qy 721 AGGTTGAGACCCCATGGCACTGCTGAGCTCGTGGGCTTACTGGAGAGACCCCGGGTA 780
Db 756 AGGTTGAGACCCCATGGCACTGCTGAGCTCGTGGGCTTACTGGAGAGACCCCGGGTA 815
Qy 781 GGGGCTGT 840
Db 816 GGGGCTGT 875
Qy 841 AGCAGCTCGCATCTGAGTACCTTCAATGTGGAGCGGCTTGTGAGACTACTTCCAC 900
Db 876 AGCAGCTCGCATCTGAGTACCTTCAATGTGGAGCGGCTTGTGAGACTACTTCCAC 935
Qy 901 TGTGTATCTGATCAGCGGCTCTAG----- 928
Db 936 TGTGTATCTGATCAGCGGCTCTAG----- 955
Qy 929 -----AATCTGCCAGGCCCCCAGGGAGACGCGATGATGCC 965
Db 996 CTTGAGGCTTGATACAAACAGAGGTTCTCGGTGATCCCACTGATCTTTTGGGGATGACCG 1055
Qy 966 CTCATTCCTCGCCC-----CGTGCAGGTACACCTCCAGGTCGCG 1006
Db 1056 CACCTCACAAACCGCATGCTCAGCATGGGTATGTACCAAGTACACCTCCAGGTCGCG 1115
Qy 1007 TGCTACTCAGAGACGCGCTCGTCTTCTCGGGTGGCTGAGCCAGACACGCTGTGTC 1066
Db 1116 TGCTACTCAGAGACGCGCTCGTCTTCTCGGGTGGCTGAGCCAGACACGCTGTGTC 1175

Qy 1067 AAGTCGTACTTCCGTGA 1083
Db 1176 AAGTCGTACTTCCGTGA 1192

RESULT 2

US-09-949-016-5659 ; Sequence 5659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5659

Query Match 85.9%; Score 929.8; DB 4; Length 2087;
Best Local Similarity 90.4%; Pred. No. 4.1e-175;
Matches 1046; Conservative 0; Mismatches 37; Indels 74; Gaps 2;

Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGCGCTGGCC 60
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGCGCTGGCC 95
Qy 61 CGGAGGCTGTGACCATGCGCTTCGCGCTGCTCATCTGCGGCTTCATGACCTGGCCCTAC 120
Db 96 CGGAGGCTGTGACCATGCGCTTCGCGCTGCTCATCTCTGCGGCTTCATGACCTGGCCCTAC 155
Qy 121 GCGCGCGGCTGCGCTGCGCTCCGATCGTACGCGCTCTGCGCTTCGCGCTCTACGCG 180
Db 156 GCGCGCGGCTGCGCTGCGCTCCGATCGTACGCGCTTCGCGCTTCGCGCTCTACGCG 215
Qy 181 GCGCTTCCTTTCAGCGCACCTGCTGCGCAGAGCCCTTCCTGCGTACCTGGAGCACCGCGCG 240
Db 216 GCGCTTCCTTTCAGCGCACCTGCTGCGCAGAGCCCTTCCTGCGTACCTGGAGCACCGCGCG 275
Qy 241 GTGGCGGCGCGCGCGCGCGCGCTGATGAGCGCACCGCGCGAGTGTGGCGCTGAC 300
Db 276 GTGGCGGCGCGCGCGCGCGCGCTGATGAGCGCACCGCGCGAGTGTGGCGCTGAC 335
Qy 301 ATCTCGGCTTACAGAGAGACCCCGCGTACCTGCGCGAGTGTGGCTGCGCGCGCGCC 360
Db 336 ATCTCGGCTTACAGAGAGACCCCGCGTACCTGCGCGAGTGTGGCTGCGCGCGCGCC 395
Qy 361 CTGCTGTATACCGCGCGCGCGCTGCGCTCTCATGCTGTGGATGGCAACCGCGCGAG 420
Db 396 CTGCTGTATACCGCGCGCGCGCTGCGCTCTCATGCTGTGGATGGCAACCGCGCGAG 455
Qy 421 GACCTCTACATGCTCGAATGCTTCGCGAGGCTTCCTGCTGAGAGAGACCCCGCGCACGTAC 480
Db 456 GACCTCTACATGCTCGAATGCTTCGCGAGGCTTCCTGCTGAGAGAGACCCCGCGCACGTAC 515
Qy 481 GTGTGGAGCGCACTACCAAGCCCTGGGAACCCCGCGCGCGCGGTGGCGCC 540
Db 516 GTGTGGAGCGCACTACCAAGCCCTGGGAACCCCGCGCGCGCGGTGGCGCC 575
Qy 541 GGAGCGCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGGCACTGGAGCGCGCTGTG 600

Db 1130 TCAGAGAGCCCTCGTCTCTCTGCTGGCTGAGCCAGACACGCTGGTCCAAGTCG 1189
QY 1073 TACTTCCGTGA 1083
Db 1190 TACTTCCGTGA 1200

RESULT 4
US-09-949-016-17401
; Sequence 17401, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17401
; LENGTH: 14862
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17401

Query Match 64.4%; Score 637.2; DB 4; Length 14862;
Best Local Similarity 99.6%; Pred. No. 5e-129;
Matches 699; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 CAGCAGGAGCGCCCAAGCCCACTCTCGAGCCCGCCGCTGCTCGGCTGGCCCGGAGG 66
Db 6073 CAGCAGGAGCGCCCAAGCCCACTCTCGAGCCCGCCGCTGCTCGGCTGGCCCGGAGG 6132
QY 67 GTGCTGACATCGCTTCCGCTGCTCATCTGGGCTCATGACCTGGGCTTACGCCGCC 126
Db 6133 GTGCTGACATCGCTTCCGCTGCTCATCTGGGCTCATGACCTGGGCTTACGCCGCC 6192
QY 127 GGGGTGCGGCTGGCTTCCGATCGCTACCGGCTCTTGGGCTTACCGGCTTACCGGCTTCC 186
Db 6193 GGGGTGCGGCTGGCTTCCGATCGCTACCGGCTCTTGGGCTTACCGGCTTACCGGCTTCC 6252
QY 187 CTTTCAGCGCACCTGGTGGCGAGAGCCTCTTGGGCTACCTGGAGCACCGGCGGGTGGCG 246
Db 6253 CTTTCAGCGCACCTGGTGGCGAGAGCCTCTTGGGCTACCTGGAGCACCGGCGGGTGGCG 6312
QY 247 GCGGCGCGCGGGGCGCTGATGACGACCGCGCGAGTGTGGGCTGACATCTCC 306
Db 6313 GCGGCGCGCGGGGCGCTGATGACGACCGCGCGAGTGTGGGCTGACATCTCC 6372
QY 307 GCCTACAGAGAGACCCCGCTGACTGCGCAGATGCTTGGGCTTCCGCGCCGCTTGGTG 366
Db 6373 GCCTACAGAGAGACCCCGCTGACTGCGCAGATGCTTGGGCTTCCGCGCCGCTTGGTG 6432
QY 367 TACCGCGCGCGGCTGGGCTTCTCATGTGGTGGATGTCACCGCGCGAGACCTC 426
Db 6433 TACCGCGCGCGGCTGGGCTTCTCATGTGGTGGATGTCACCGCGCGAGACCTC 6492
QY 427 TACATGTCGACATGTTCCGCGAGGTCTTCCGTCAGCAGAGACCCCGCACCTAGCTGTGG 486
Db 6493 TACATGTCGACATGTTCCGCGAGGTCTTCCGTCAGCAGAGACCCCGCACCTAGCTGTGG 5552
QY 487 GACGCGAATACCAACAGCCCTGGGAAACCGCGCGCGCGGCGGCTGGGCGCGGAGCC 546
Db 6553 GACGCGAATACCAACAGCCCTGGGAAACCGCGCGCGCGGCGGCTGGGCGCGGAGCC 6612

QY 547 TATCGGAGGTGGAGCGGAGATCCCTGGCGGCTGGCAGTGGAGCGCTGGTGAAGACT 606
Db 6613 TATCGGAGGTGGAGCGGAGATCCCTGGCGGCTGGCAGTGGAGCGCTGGTGAAGACT 6672
QY 607 GCGAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCTATGTACAGCC 666
Db 6673 GCGAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCTATGTACAGCC 6732
QY 667 TTCAAGCGCTCGAGAGATTCGCTGACCTACGTGAGGTCTGT 708
Db 6733 TTCAAGCGCTCGAGAGATTCGCTGAGCTACGTGAGGTGAGT 6774

RESULT 5

US-09-949-016-201775/c
; Sequence 201775, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201775
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201775

Query Match 36.6%; Score 396.4; DB 4; Length 601;
Best Local Similarity 99.2%; Pred. No. 1.2e-69;
Matches 397; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 309 CTACAGAGAGACCCCGGTACCTGCGCAGTGTGCGTCCGCGCGCCCTGCTCTA 368
Db 601 CTACAGAGAGACCCCGGTACCTGCGCAGTGTGCGTCCGCGCGCCCTGCTCTA 542
QY 369 CCCGCGCGCGGCTGGCGCTCTCATGTGTGGATGGCAACCGCGCGAGGACCTCTA 428
Db 541 CCCGCGCGCGGCTGGCGCTCTCATGTGTGGATGGCAACCGCGCGAGGACCTCTA 482
QY 429 CATGCTGACATGTTCCGCGAGGTCTTCCGTGACGAGAACCCCGCACGTACGTGTGGGA 488
Db 481 CATGCTGACATGTTCCGCGAGGTCTTCCGTGACGAGAACCCCGCACGTACGTGTGGGA 422
QY 489 CGGCAACTACCAACAGCCCTGGGAAACCCCGCGCGCGCGCTGGGCGCGGAGCCTA 548
Db 421 CGGCAACTACCAACAGCCCTGGGAAACCCCGCGCGCGCGCTGGGCGCGGAGCCTA 362
QY 549 TCGGAGGTGAGAGCGGAGGATCTTGGCGGCTTCCGCTGAGAGCGCTGTGTGAGACTCG 608
Db 361 TCGGAGGTGAGAGCGGAGGATCTTGGCGGCTTCCGCTGAGAGCGCTGTGTGAGACTCG 302
QY 609 CAGTGTGCTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGGAGGTCTATGTACAGCCCTT 668
Db 301 YAGTGTGCTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGGAGGTCTATGTACAGCCCTT 242
QY 669 CAAGCGCTCGGAGATTCCGTGGAGTACGTGAGGTCTGT 708
Db 241 CAAGCGCTCGGAGATTCCGTGGAGTACGTGAGGTGAGT 202

RESULT 6

US-09-949-016-3728

; Sequence 3728, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3728
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-3728

Query Match 25.2%; Score 273; DB 4; Length 1662;
Best Local Similarity 61.1%; Pred. No. 3.3e-45;
Matches 532; Conservative 0; Mismatches 300; Indels 39; Gaps 4;
Qy 61 CGGAGGGGTGACCATCGCTTCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 28 CGTGTGTGGGACACGAGCTGTGGCTTGGGCTGCTGGGCTGCTGGGCTGCTAT 87
Qy 121 GCCCGCGGGGTGCGCTGCGCTGCGATCGCTACGCGCTCTGGGCTTGGGCTTACGGG 180
Db 88 GTGACGGGCTACCACTTATCATCACAGGAAAGCACTACCTGTCTCGGCTGTACGGC 147
Qy 181 GCCCTCTCTTTCAGGCGACCTGGTGGCGAGAGCCTTTCGGGTACCTGGAGACCGGGCG 240
Db 148 GCCATCTCTGGGCTGACCTGCTATTTCAGAGCCTTTTGGCTTCTGGAGCACCGGGCG 207
Qy 241 GTGCGCGCGGGCGCGGGCGCTGATGACGACCGCGC---GCGGTGTGGCGCTG 297
Db 208 ATGCGAGCTGCGGGCGAGCCTTGAAGCTGCTCCCGCGGGGCTCGGTGGCACTG 267
Qy 298 ACCATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGCCTGGCGTCCGCGCGC 357
Db 268 TGCATTGCGCATACAGAGAGACCCCTGACTACTTGGCGAAGTGCTGCGCTCGGCCAG 327
Qy 358 GCCCTGCTGTACCGCGCGGGCTGCGCTTCTCATGTGTGTGGATGGCAACCGCGCC 417
Db 328 CGCATCTCTCTCCCTGAC-----CTCAAGGTGGTTCATGTTGGTGGATGGCAACCGCCAG 381
Qy 418 GAGGACCTCTACATGTTGCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCAG 477
Db 382 GAGAGCGCTACATGCTGACATCTTCACAGAGTGTGGCGGACCGAGACGCGCGC 441
Qy 478 TACGTGTGGGACGGCACTACACAGCCCTGGGAACCCCGCGGGCGCGCGCTGGGC 537
Db 442 TCTTTGTGTGCGCAGCACTTCA-----TGAGGACAGGCGAGGTGAG 486
Qy 538 GCGGAGCTATCGGAGGTGAGCGGAGGATCTTGGGCGGCTGGCAGTGGAGCGCGCTG 597
Db 487 ACGGAGGCGACCTGTCAGGAGGATGACCGT-----GTGCGGGATGTG 531
Qy 598 GTGAGGACTCGCAGGTGTGTGGTGGCGAGCGCTGGGCGGCGCAAGCGGAGGTGATG 657
Db 532 GTGCGGGCGACACCTTCTCGTGATCATGTCAGAAAGTGGGAGGCAAGCGGAGGTGATG 591
Qy 658 TACAGCGCTTCAAGCGGCTCGGAGATTCGGTGGACTACGTGCGAGGTCTGTGACTCGGAC 717
Db 592 TACAGCGCTTCAAGCGGCTCGGAGATTCGGTGGACTACATCCAGGTGTGCGACTCTGAC 651
Qy 718 ACAAGGTTGGAACCCATGGCACTGCTGTGAGCTCTGTGGGCTGTGAGAGGACCCCGG 777
Db 718 ACAAGGTTGGAACCCATGGCACTGCTGTGAGCTCTGTGGGCTGTGAGAGGACCCCGG 777

Db 652 ACTGTGCTGGATCCAGCTGCAACCATCGAGATGTTTCGAGTCTTGGAGGAGGATCCCA 711
Qy 778 GTAGGGGCTGTGTGGGAGCTGCGGATCTTAACCCCTCTGGACTCTCTGGGTGAGCTTC 837
Db 712 GTAGGGGAGTCTGGGGAGATGTCAGATCTCTCAAGTACGACTCATGATTCCTTC 771
Qy 838 CTAAGCAGCTGCGATCTGCGGTAGCTTCAATGTGAGAGCGGCTTGTTCAGAGCTACTTC 897
Db 772 CTGAGCAGCTGCGGTACTGATGCGCTTCAACGTGGAGCGGCTGCCAGTCTCTACTTT 831
Qy 898 CACTGTGTATCTGATCATCAGCGGTCTCTAG 928
Db 832 GCGTGTGTGAGTGTATTTAGTGGGCGCTTGG 862

RESULT 7

US-08-812-008-31
; Sequence 31, Application US/08812008
; Patent No. 6602893
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustine, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1662
; OTHER INFORMATION:
US-08-812-008-31

Query Match 21.8%; Score 236.4; DB 4; Length 1665;
Best Local Similarity 58.5%; Pred. No. 5.7e-38;
Matches 511; Conservative 0; Mismatches 321; Indels 42; Gaps 4;
Qy 61 CGGAGGGTGTGACCATCGGCTTGGCCCTGCTCATCTCTGGCCTCATGACCTGGGCTTAC 120
Db 28 CGTGTGTGGGACCCAGTCTGTTTGGCCCTGTTGTGCTGGAGGACATCTTGGCGGCTTAT 87

```
Qy 121 GCCGCGGGTGGCGCTGGCTCCGATCGCTACGCGCTCTGGCTCTGGCTCTACGGG 180
Db 88 GTGACAGGCTACCAAGTTTATCCACAGAAAGACATACCTGTCTTTGGGCTCTACGGT 147
Qy 181 GCCTTCTCTTTTCAGCGCACCTGGTGGCGCAGAGCTCTTCGCGTACCTCGAGACACCGCGG 240
Db 148 GCCATCTCTGGTCTACATCTGCTCATCCAGAGCTCTTTGGCTTCTTGAGACACCGTCA 207
Qy 241 GTGGGGGGGGGGGGCG-----GGGGCGCTGATGAGCCACCGCGCGAGTGTGGCG 294
Db 208 ATGGCGAGGGCAGGGCGCGCCCTCAAGCTGCACTCTCCAGAGAGTGGCTTCAGTGCGCA 267
Qy 295 CTGACCATCTCCGCTACACAGAGAGACCCCGCTACCTGCGCCAGTGCCTGGCTCGGCC 354
Db 268 CTCTGCAATGCTGCTTACCAAGAGAGACCCCGAATACCTGCGCAAGTGCCTTCGCTCAGCT 327
Qy 355 CGCGCCCTCTGTATACCGCGCGCGCTGGCGCTCTCATGCTGGTGGATGGCAACCGC 414
Db 328 CAGCGCATTTCTTCCAAAC-----CTCAAGGTGGTCACTGATGGATGGCAATCGC 381
Qy 415 GCCAGGACCTCTACATGCTGCAATGTTCCGCGAGGTTCTTCGCTGAGAGGACCCCGCC 474
Db 382 CAGGAAGATACCTACATGTTGGACATCTTCCATGAGGTGCTGGGTGGCACTGAGCAAGCT 441
Qy 475 ACGTACGTGTGGGACGGCACTACACAGCCCTGGGNAACCGCGGGGGCGCGGTG 534
Db 442 GGCTTCTTTGTGGGTGAGCAATTTCCAT-----GAGGCGGGTGAAGGA 486
Qy 535 GCGCGCGAGCTATCGGAGGTGGAGCGGAGGATCTGGCGCGCTGGCAGTGGAGGCG 594
Db 487 GAGACAGAGCGCCTGCAGAGGATGAGCGT-----GTGCGAGCT 531
Qy 595 CTGTGAGGACTCGAGGTGCTGTGCTGCGCGCAGCGCTGGCGCGGCAAGCGCGAGGTC 654
Db 532 GTGCTGTGGCGCAGCACCTTCTCATGTCATCATGAGAAAGTGGGGGGGCAAGCGTGAGGTC 591
Qy 655 ATGTACACAGCTTCAGGCGCTCGAGATTCGGTGACCTACGTCAGGTCGTGACTCG 714
Db 592 ATGTACACTGCTTCAGAGGCCCTTGGCAACTCAGTGACATACATCCAGGTTGTGACTCT 651
Qy 715 GACACAAAGTTGGACCCCATGCGACTGTGTGAGCTCTGTGGGCTCTGGAGAGAGCC 774
Db 652 GACACTGTCTGGACCCAGCCCTGCACCATTGAGATGCTTCGAGTCTTGGAGAGATCCC 711
Qy 775 CGGTAGGGCTGTGTGGGAGCTGGGATCTTAACTCTGAGATCTCTGGGTCAGC 834
Db 712 CAAGTAGGAGGTGTGGAGGAGATGTCCAAATCTCAACAGTATGATTCATGGATCTCC 771
Qy 835 TTCTTAAGCAGCCCTGCATCTGGGTAGCCCTTCAATGTGGAGCGGCTTGTGAGGATC 894
Db 772 TTCTGAGCAGTGTGAGGTACTGGATGCTTCAACGTGGAGCGGGCTGCGAGTCTCTAC 831
Qy 895 TTCCACTGTGTATCTCTGCATCAGCGGCTCTCTAG 928
Db 832 TTTGGCTGTGCAATGATTAAGTGGGCTTTGG 865

RESULT 8
US-09-949-016-201780/c
; Sequence 201780, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201780
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201780

Query Match 21.2%; Score 229.4; DB 4; Length 601;
Best Local Similarity 98.7%; Pred. No. 1.3e-36;
Matches 230; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 696 CGTGCAGTCTGTGACTCGGACACAGGTTGGACCCCATGACACTGCTGGAGCTCGTGCG 755
Db 403 CTTACAGGCTGTGACTCGGACACAGGTTGGACCCCATGACACTGCTGGAGCTCGTGCG 344
Qy 756 GGTACTCGACGAGGACCCCGGGTAGGGGCTGTGGTGGGGACGTGGCGGATCCCTTAACCC 815
Db 343 GGTACTCGACGAGGACCCCGGGTAGGGGCTGTGGTGGGGAYGTGGGATCCCTTAACCC 284
Qy 816 TCTGAGCTCTCGGTGAGCTTCTTAAGCAGCTCGGATCTGGGTAGCCTTCAATGTGA 875
Db 283 TCTGAGCTCTCGGTGAGCTTCTTAAGCAGCTCGGATCTGGGTAGCCTTCAATGTGA 224
Qy 876 GCGGGCTTGTGAGAGCTTCTTCCACTGTGTATCTGTCATCAGCGGTCTCTAG 928
Db 223 GCGGGCTTGTGAGAGCTTCTTCCACTGTGTATCTGTCATCAGCGGTCTCTAG 171

RESULT 9
US-08-675-499A-1
; Sequence 1, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2947 base pairs
; TYPE: nucleic acid
```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-675-499A-1

Query Match 16.6%; Score 179.4; DB 4; Length 2947;
Best Local Similarity 55.3%; Pred. No. 1.1e-26;
Matches 441; Conservative 0; Mismatches 311; Indels 45; Gaps 3;

Qy 159 CTGGCTTGGGCTTACGGGGCTTCTTTCAGCGCACCTGGTGGCGCAGAGCTCTT 218
Db |||||
Qy 636 CTTCTCATTTGGACTGAGTGGCTTTTACCTCGCATCTCATCATCAAAGCTCTT 695
Db |||||
Qy 219 CGCGTACCTGAGCACCAGCGGGTGGCGCGCGCGGGCGCGCTGGATGAGCCAC 278
Db |||||
Qy 696 TGCCTTTTGGAAACACCGGAAATGAAGAAGTCCCTTGAACCCCGATTAAATGAACAA 755
Db |||||
Qy 279 CGCGCGAGTGTGGGCTGACATCTCCGCTTACAGGAGACCCCGGTACCTGCGCA 338
Db |||||
Qy 756 AACG-----GTAGCACTCTGCATCGCTGCGTACCAAGAGACCTGACTACTACGGA 809
Db |||||
Qy 339 GTGCTGCGCTCGCGCGCGCTGCTGCTACCGCGCGCGCGCTGCGCTCCTCATGCT 398
Db |||||
Qy 810 ATGTTTGCATCTGTGAAGAGCTGACCTACCTCTG-----GGATTAAAGTCGTGATGGT 863
Db |||||
Qy 399 GGTGGATGGCAACCGCGCGCGAGGACCTCTCATGCTGACATGTTCCGCGAGGCTCTCGC 458
Db |||||
Qy 864 CATCGATGGGAACCTCAGACGACGACCTTTATGATGACATATTCAGCGAAGTATTGG 923
Db |||||
Qy 459 TGACGAGAACCCCGCCACGCTACGCTGTGGGACGGCAACTACCAAGCCCTGGGAACCCGC 518
Db |||||
Qy 924 CAGGACAAATCGGCGACGCTACATCTGGAAGAACTTTCATGAAAG----- 972
Db |||||
Qy 519 GCGCGCGCGCGGTGGCGCGCGAGCTATCGGAGGTGGAGCGGAGGATCCTGGCG 578
Db |||||
Qy 973 -----GGACCTGGTGAGACAGAAAGATCCCATAAAGAAAGTTC 1010
Db |||||
Qy 579 GCTGGCAGTGAGCGGCTGGTGGAGTCTCGCAGGTGGTGGCGCGCGAGCGCTGGG 638
Db |||||
Qy 1011 ACAACATGTCACCCAAATTTGCTTGTCTTAAACAAAGTATTTGCATCATGCAAAATGGG 1070
Db |||||
Qy 639 CGCAAGCGCAGGTCATGTACACAGCTTCAAGCGCTCGGAGATTCGGTGGAGTACGT 698
Db |||||
Qy 1071 TGAAGAGAGAAGTCTATGTACACAGCTTCAGAGCACTGGGCGGAAGCGTGGATTATGT 1130
Db |||||
Qy 699 GCAGTCTGACCTCGGACACAGTTGGACCCCATGCTGCTGGAGCTCGTGGGCT 758
Db |||||
Qy 1131 ACAGGTGTGACTCAGATATATGCTTGAACCTGCTCATCTGTGGAGATGGTGAAGGT 1190
Db |||||
Qy 759 ACTGACGAGAACCCCGGCTAGGGCTGTGGTGGGACGCTGGGATCTCTTAACCTCT 818
Db |||||
Qy 1191 CTTAGAGGAAGCCCTATGTTGGAGGTGTTGGAGGATGTCAGATTTTAAACAGTA 1250
Db |||||
Qy 819 GACTCTGGGTGAGCTTCTTAAGCAGCTCGATACCTGGTGGAGCTTCAATGTGAGCG 878
Db |||||
Qy 1251 TGATTCCTGGATCTCTTCTCAGCAGCGTGAGATCTGGATGGCTTTTAAATATAGAAAG 1310
Db |||||
Qy 879 GGCTGTGAGAGTACTTCCACTGTATCTGATCAGCGGTCTCTAGATCCTGCCC 938
Db |||||
Qy 1311 GGCCTGCCAGTCTTATTTTGGCTGTCCAGTGCAAGAGCGGTCTCTGGGAATGTACAG 1370
Db |||||
Qy 939 AGGCCAGGAGGACG 955
Db |||||
Qy 1371 AAATCTTGTGCAATG 1387
Db |||||

RESULT 10

US-08-812-008-1
; Sequence 1, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.

APPLICANT: Augustine, M. L.
TITLE OF INVENTION: GENE ENCODING HYALURONAN
TITLE OF INVENTION: SYNTHASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
City: Minneapolis
STATE: MN
COUNTRY: U.S.A
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,008
FILING DATE: 05-MAR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/675,499
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Embretson, Janet E
REGISTRATION NUMBER: 39,665
REFERENCE/DOCKET NUMBER: 150.183US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2947 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-812-008-1

Query Match 16.6%; Score 179.4; DB 4; Length 2947;

Best Local Similarity 55.3%; Pred. No. 1.1e-26;
Matches 441; Conservative 0; Mismatches 311; Indels 45; Gaps 3;

Qy 159 CTGGCTTGGGCTTACGGGGCTTCTTTCAGCGCACCTGGTGGCGCAGAGCTCTT 218
Db |||||
Qy 636 CTTCTCATTTGGACTGAGTGGCTTTTACCTCGCATCTCATCATCAAAGCTCTT 695
Db |||||
Qy 219 CGCGTACCTGAGCACCAGCGGGTGGCGCGCGCGGGCGCGCTGGATGAGCCAC 278
Db |||||
Qy 696 TGCCTTTTGGAAACACCGGAAATGAAGAAGTCCCTTGAACCCCGATTAAATGAACAA 755
Db |||||
Qy 279 CGCGCGAGTGTGGGCTGACCATCTCGGCTTACAGGAGACCCCGGTACCTGCGCA 338
Db |||||
Qy 756 AACG-----GTAGCACTCTGCATCGCTGCGTACCAAGAGACCTGACTACTACGGA 809
Db |||||
Qy 339 GTGCTGCGCTCGCGCGCGCTGCTGCTACCGCGCGCGCGCTGCGCTCCTCATGCT 398
Db |||||
Qy 810 ATGTTTGCATCTGTGAAGAGCTGACCTACCTCTG-----GGATTAAAGTCGTGATGGT 863
Db |||||
Qy 399 GGTGGATGGCAACCGCGCGCGAGGACCTCTCATGCTGACATGTTCCGCGAGGCTCTCGC 458
Db |||||
Qy 864 CATCGATGGGAACCTCAGACGACGACCTTTATGATGACATATTCAGCGAAGTATTGG 923
Db |||||
Qy 459 TGACGAGAACCCCGCCACGCTACGCTGTGGGACGGCAACTACCAAGCCCTGGGAACCCGC 518
Db |||||
Qy 924 CAGGACAAATCGGCGACGCTACATCTGGAAGAACTTTCATGAAAG----- 972
Db |||||
Qy 519 GCGCGCGCGCGGTGGCGCGCGAGCTATCGGAGGTGGAGCGGAGGATCCTGGCG 578
Db |||||
Qy 973 -----GGACCTGGTGAGACAGAAAGATCCCATAAAGAAAGTTC 1010
Db |||||
Qy 579 GCTGGCAGTGAGCGGCTGGTGGAGTCTCGCAGGTGGTGGCGCGCGAGCGCTGGG 638
Db |||||

Db 1011 ACAACATGTACCCAAATTGGTCTTGTCTAACAAAAGTATTTGCATCATGCAAAAATGGG 1070
Qy 639 CGGCAAGCGCAGGTGATGATACAGAGCTTCAAGCGCGTCGGAGATTCGGTGGACTACGT 698
Db 1071 TGGAAAGAGAGAAGTATGATGATACAGAGCTTCAAGAGCACTGGGGCGAAGCGTGGATTATGT 1130
Qy 699 GCAGGTCTGTACCTCGGACACAAAGTTGGACCCCATGGCACTGCTGGAGCTCGTGGCGGT 758
Db 1131 ACAGGTGTGACTCAGATATATGCTTGACCCCTGCCTCATCTGTGGAGATGGTGAAGGT 1190
Qy 759 ACTGACAGAGACCCCGGTAGGGCTGTGTGGGAGCGTGGCGATCCTTAAACCCCTCT 818
Db 1191 CTTAGAGAGACCCCTATGTTGGAGGTGTGGAGGATGTCAGATTTTAAACAAGTA 1250
Qy 819 GGAATCTCGGTGAGCTTCTTAAGACGCTCGGATACCTGGTAGCCTTCAATGTGGAGCG 878
Db 1251 TGATTCCTGGATCTCTTCTCAGAGCGTGGAGATCTGGATGGCTTTAATATAGAAAG 1310
Qy 879 GGCTGTGAGAGCTACTTCCACTGTGTATCTGATCAGCGGTCTCTAGAACTCTGCC 938
Db 1311 GGCTGTGAGCTTATTTTGGCTGTGTCCAGTGAAGCGGTCTCTGGGAATGTACAG 1370
Qy 939 AGGCCCCAGGAGCACG 955
Db 1371 AAATCTTGTGTCATG 1387

RESULT 11
US-09-949-016-15470
; Sequence 15470, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15470
; LENGTH: 9871
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15470

Query Match 15.5%; Score 167.6; DB 4; Length 9871;
Best Local Similarity 58.9%; Pred. No. 2.7e-24;
Matches 384; Conservative 0; Mismatches 229; Indels 39; Gaps 4;
Qy 61 CGGAGGTGTGACCATGGCTTCGCCCTGCTCATCTCGGCTCATGAGCTGAGCTGGGCTTAC 120
Db 2028 CGTGTGTGGGCACACAGCTGTTTGGCTGCTGAGTGTGCTGGTGGCTCTGGCAGCTAT 2087
Qy 121 GCCCGCGGGTGGCGCTGGCTCGATCGCTACGGCTCTGGGCTTGGGCTTACGGG 180
Db 2088 GTGACGGGCTACCAAGTTCATCCACGGAAGCACTACCTGTCTTGGGCTGTACGGC 2147
Qy 181 GCCTTCTCTTCAGGCGACCTGTTGGCGCAGAGCTCTTTCGGTACCTGGAGCACCGCGG 240
Db 2148 GCCATCTTGGGCTGACCTGCTCATTCAGAGCTTTTTCGCTTCTTGGAGCACCGGCGC 2207
Qy 241 GTGCGGCGGCGCGCGGGGCGCTGTGATGACGCCACCGGC---GAGTGTGGCGCTG 297
Db 2208 ATGCGAGTGTGCGGCCAGGCGCTGAAGTGTGCTTCCCGCGCGGGGCTGGTGGCACTG 2267

Qy 298 ACCATCTCCGCTACCAGGAGGCCCGCGCTACTCTGGCGCAGTGCCTGGCGCTCCGCCGC 357
Db 2268 TGCATTCGCGCATACAGAGGAGCCCTGACTACTTGGCGAAGTGCCTGGCTCGGCCAG 2327
Qy 358 GCCTTGTGTACCGCGCGCGCTGGCTCCTCATGTGTGGATGGCAACCGCGCC 417
Db 2328 CGCATCTCCTTCCCTGAC-----CTCAAGGTGTGTGATGGTGGCAACCGCCAG 2381
Qy 418 GAGGACCTTACATGTGTCGACATGTTCCGGAGGTCTTCTGCTGACGAGGACCCCGCCACG 477
Db 2382 GAGGACCCCTACATGCTGGACATCTTCCACAGGTGCTGGCGCGCACCGAGACAGCGCGC 2441
Qy 478 TACGTGTGGGACGCAACTACACAGCCCTGGGAACCCCGCGCGCGCGCGCGCTGGGC 537
Db 2442 TTCTTGTGTGGCGACCAACTTCCAT-----GAGGACGCGAGGGTGAG 2486
Qy 538 GCCGAGCCCTATCGGAGGTGGAGCGGAGGATCCTGGCGGCTGGCAGTGGAGCGCGTG 597
Db 2487 ACGAGGCCAGCCCTGCGAGGAGGATGACCGT-----GTGCGGATGTG 2531
Qy 598 GTGAGGACTCGAGGTGCTGTGGTGGCGCAGCGCTGGGGCGCAAGCGGAGTCAATG 657
Db 2532 GTGCGGCCAGCACCTTCTGTCATCATCAGAAAGTGGGGAGGCAAGCGGAGTCAATG 2591
Qy 658 TACACAGCCTTCAAGCGCTCGGAGATTCTGCTGACTACGTGCAGGTCTGTG 709
Db 2592 TACAGCCCTTCAAGGCCCTCGCGATTCTGGTGGACTACATCCAGGTAAGGG 2643

RESULT 12

US-09-949-016-737
; Sequence 737, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 737
; LENGTH: 3003
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-737

Query Match 14.2%; Score 153.6; DB 4; Length 3003;
Best Local Similarity 53.2%; Pred. No. 1.4e-21;
Matches 426; Conservative 0; Mismatches 329; Indels 45; Gaps 3;
Qy 164 CCTTGGCCCTTACGGGCGCTTCTTTACGCGCACCTGGTGGCGCAGAGCTTTCGCGT 223
Db 669 CTTTGGACTGTATGTGTGCTTTTGGCATCACACCTCATCATCAAAGCTGTGTGCT 728
Qy 224 ACCTGAGACCGCGGGTGGCGGGCGGGGGCGCTGGATGACGCCACCGCGC 283
Db 729 TTTTGGAGCACCCGAAAAATGAAAAATCCCTAGAAACCCCAATAAGTTGAACA----- 782
Qy 284 GCAGTGTGGCGCTGACCATCTCGCCTTACCAGAGGACCCCGGTACTCTGCCAGTGGC 343
Db 783 AACAGTGTGCTTGTGATCGCTTATCAAGAGATCCAGACTACTTAAAGGAATGTT 842
Qy 344 TGGCGTCCGCCCGCGCCCTGCTGTACCCGCGCGCGGGCTGCGGCTCCTCATGTGTGTGG 403
Db 843 TGCAATCTGTGAAAGGCTAACCTACC-----CTGGGATTAAGTTGTCTATGTCATAG 896

Qy	404	ATGGCAACCGCGCGAGGAACTCTACATGTCGACATGTTCCGCGAGGTCTTCGCTGACG	463
Db	897	ATGGGAATCAGAAGATGACCTTTACATGATGGACATCTTCAGTGAAGTCATGGCGAGAG	956
Qy	464	AGGACCCCGCCACCTACGTGTGGAGACGGCACTACCAACGACCCCTGGGAACCCGCGCGG	523
Db	957	ACAAATCAGCCACCTTATATCTGGAGAACAACATTTCCACGAAAG-----	1000
Qy	524	CGGCGCGGTGGGCGCGGAGCCTTATCGGAGGTGGAGGGCGGAGGATCCTGGGCGCGCTGG	583
Db	1001	-----GGTCCCCTGGAGACAGATGAGTCACTAAAGAAAGCTCGCAAC	1043
Qy	584	CAGTGGAGCGCTGGTGAGGAACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCA	643
Db	1044	ACGTAAACCAATTGGTCTTGTCACAAAGATATCTGCATCATGCAAAAATGGGTGGAA	1103
Qy	644	AGCGGAGGTCAATGACACAGCCTTCAAGGCGCTCGAGATTCGGTGGCACTACGTGCAGG	703
Db	1104	AAAGAGAAGTCATGTCACACAGCCTTACAGAGCACTGGGACGAAGTGTGGATTATGTACAGG	1163
Qy	704	TCTGTGACTGGGACACAAGGTTGGAACCCCATGGCACTGTCTGGAGCTCGTGGGGTACTGG	763
Db	1164	TTTGTGATTGAGACACTATGCTTGACCAGCCTCATCTGTGGAGATGGTAAAAAGTTTTAG	1223
Qy	764	ACGAGGACCCCGGTAGGGCTGTTGGTGGGAGCGTGGGATCCTTAAACCTCTGGACT	823
Db	1224	AAGAAGATCCCATGCTGTGGAGGTGTGGGGAGATGTCCAGATTTTAAAAAAGTACGATT	1283
Qy	824	CCTGGGTGAGTCTTCAAGCAGCCTGCGATACGGGTAGCCTCAATGTGGAGCGGGCTT	883
Db	1284	CCTGGATCTCAATCCTCAGCAGTGTAAGATATTGGATGGCTTTTAAATATAGAAGGGCCT	1343
Qy	884	GTCAGAGTACTTCCACTGTGTATCTTCGCATCAGCGTCTCTAGAAATCTCGCCAGGCC	943
Db	1344	GTCACTCTATTTTGGGTGTGTTTCAGTGCATTAGTGGACCTCTGGGAATGTACAGAAACT	1403
Qy	944	CCAGGGAGCAGCGCATGATG	963
Db	1404	CTTTGTTGCATGATTTGTG	1423

```

RESULT 13
US-09-949-016-132199
; Sequence 132199, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J, Craig et al.
; TITLE OF INVENTION: POLYMERISMS IN KNOWN GENES ASSOCIATED
; WITH HUMANS WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132199
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-132199

Query Match      12.7%; Score 137.4; DB 4; Length 601;
Best Local Similarity 58.5%; Pred.No. 2e-18;
Matches 300; Conservative 0; Mismatches 201; Indels 12; Gaps 3;

          61 CGGAGGGTGCTGCCATCGCCCTTCGCCCTGTCTCATCTGGGGCCTCATGCCTGGGCCTAC 120
          ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db	50	CGTGTGTGGGACACAGCCTGTTTCCCTGGCAGTGTGGGTGGCATCTCTGGCAGGCTAT	109
Qy	121	GCCECGGGGGTGCCTGGCTCCCGATCGCTACGGCTCTCTGGCCCTTACGGG	180
Db	110	GTGAGGGGTACACAGTTTCATCCACAGGAAAGCACCTACCTGTCTTTCGGCCCTGTACGGC	169
Qy	181	GCCTTCCTTTACGGGCACCTGGTGGCGCAGAGCCTCTTCCGCTACCTGGAGCACCGCGG	240
Db	170	GCCATCCTTGGGCTGGCACCTGCTCATTTACAGAGCCTTTTTCCTTCCCTGGAGCACCGGCGC	229
Qy	241	GTGGCGGCGCGCGCGGGGGCGCTGGATGTCAGCCACACCGCGC---GCAGTGTGGCGCTG	297
Db	230	ATGCGAGCTGCGCGGCAGGCCCTGAAGCTGCCCTCCCGCGGGGGCTCGTGGGCACTG	289
Qy	298	ACCATCTCCGCCTACACAGAGAGACCCCGCGTACTCTGCGCAGTGCCTTGGCGTCCGCCCGC	357
Db	290	TGCATTGCCGCTACACAGAGAGACCCCTGACTACTTTCGCAAGTGCCTTGCCTCGGCCAG	349
Qy	358	GCCTGTGTACCCGCGCGCGGCTGCGGTCCTCATGTGTGTGGATGGCAACCGCGCC	417
Db	350	CGCATCTCTTCCTCTGAC-----CTCAAGTGGTCATGGTGTGGATGGCAACCGCGCAG	403
Qy	418	GAGGACCTCTACATGTGTCGACATGTTCCGCGAGGCTTT---CCCTGACAGGAGACCCGCGC	474
Db	404	GAGGACGCTACATGCTGGACATTTCCACAGAGTGTCTGGCGCGCACCGAGCAGGCGCGC	463
Qy	475	ACGTACGTGTGGACGGCAACTACCAACAGCCTCTGGGAAACCCCGCGCGCGCGCGGTG	534
Db	464	TTCTTTGTGTGGCGCAGCAACTTCCATGAGCGAGGCGAGGTGAGACGAGGCGCAGCCTG	523
Qy	535	GGCGCGGAGCCTATCGGAGGTGGAGCGGAG	567
Db	524	CAGAGGGCATGACCGGTGTGCGGGATGTGGT	556
RESULT 14			
US-09-902-540-8819			
; Sequence 8819, Application US/09902540			
; Patent No. 6833447			
; GENERAL INFORMATION:			
; APPLICANT: Goldman, Barry S.			
; APPLICANT: Hinkle, Gregory J.			
; APPLICANT: Slater, Steven C.			
; APPLICANT: Wiegand, Roger C.			
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof			
; FILE REFERENCE: 38-10(15849)B			
; CURRENT APPLICATION NUMBER: US/09/902,540			
; CURRENT FILING DATE: 2001-07-10			
; PRIOR APPLICATION NUMBER: 60/217,883			
; PRIOR FILING DATE: 2000-07-10			
; NUMBER OF SEQ ID NOS: 16825			
; SEQ ID NO 8819			
; LENGTH: 6645			
; TYPE: DNA			
; ORGANISM: Myxococcus xanthus			
US-09-902-540-8819			
Query Match 7.6%; Score 82.8; DB 4; Length 6645;			
Best Local Similarity 44.8%; Pred. No. 1.6e-07;			
Matches 357; Conservative 0; Mismatches 437; Indels 2; Gaps 1;			
Qy	3	GAGACAGCAGGACGCGCCCAAGCCCACTCTGTCAGACCCGCGCTGCTCCGGCTGGCCCG	62
Db	3823	GAGACGCCGCCACCGCCGAGGTGCTTACCGCGCGCGCGGAGCACTCTTCGACGACCCG	3882
Qy	63	GAGGGTGTGACCATCGCC--TTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC	120
Db	3883	AAGACGCCCATACACGCGCTGTTGCCGTGGCGCGCTGCGTCCCGGATGCTCGTC	3942
Qy	121	GCGCGGGGTGCGCTGGCTCCGATCGGTACGCGCTCTTGGCTCTTGGCTCTACGGG	180
Db	3943	ATCGACGGGCGCTGGCGGGGCTGCACGCGCTGGGACGTATGCGGACCTGCTGGACGTG	4002

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 08:01:42 ; Search time 762.346 Seconds
(without alignments)
8450.560 Million cell updates/sec

Title: US-10-672-399-5

Perfect score: 1083

Sequence: 1 atgagacagcagcgccgcccc.....tccaagtgtactccgtga 1083

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 297463231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:**
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1083	100.0	1083	18	US-10-672-399-5
2	953.4	88.0	1071	18	US-10-672-399-3
3	933	86.1	1737	18	US-10-672-399-1
4	933	86.1	2116	13	US-10-042-523-1
5	928	85.7	1065	18	US-10-672-399-7
6	677.8	62.6	1752	10	US-09-902-939-1
7	422	39.0	662	18	US-10-363-345A-13975
8	422	39.0	662	18	US-10-363-345A-13976
9	393.8	36.4	662	18	US-10-363-345A-13973
10	393.8	36.4	662	18	US-10-363-345A-13974
11	339.2	31.3	490	10	US-09-918-995-27210

12	273	25.2	1662	17	US-10-295-027-369	Sequence 369, App
13	273	25.2	1662	17	US-10-188-832-136	Sequence 136, App
14	250	23.1	1767	17	US-10-309-560-9	Sequence 9, Appli
15	236.4	21.8	1665	10	US-09-902-939-3	Sequence 3, Appli
16	179.4	16.6	1659	10	US-09-902-939-2	Sequence 2, Appli
17	179.4	16.6	4194	15	US-10-262-526-3	Sequence 3, Appli
18	153.6	14.2	3003	10	US-09-918-624B-64	Sequence 64, Appli
19	153.6	14.2	3003	15	US-10-262-526-1	Sequence 1, Appli
20	153.6	14.2	3387	15	US-10-084-1817-184	Sequence 184, App
21	153.6	14.2	3387	16	US-10-252-157-84	Sequence 84, Appli
22	153.6	14.2	4018	10	US-09-814-353-19942	Sequence 13942, A
23	100.2	9.3	793	9	US-09-910-943-552	Sequence 552, App
24	72	6.6	15559	18	US-10-646-664-1	Sequence 1, Appli
25	71	6.6	2182	18	US-10-437-963-62269	Sequence 62269, A
26	70	6.5	43058	9	US-09-954-456-292	Sequence 292, App
27	70	6.5	43058	9	US-09-954-456-529	Sequence 529, App
28	70	6.5	43058	9	US-09-880-107-3950	Sequence 3950, Ap
29	69.4	6.4	5784	14	US-10-152-886-64	Sequence 64, Appli
30	68.6	6.3	1755	18	US-10-437-963-78301	Sequence 78301, A
31	68.4	6.3	2214	15	US-10-156-761-2510	Sequence 2510, Ap
32	68.4	6.3	9025608	15	US-10-156-761-1	Sequence 1, Appli
33	67.6	6.2	1380	18	US-10-425-115-170699	Sequence 170699,
34	66.6	6.1	3957	16	US-10-300-562-193	Sequence 193, App
35	66.6	6.1	3957	16	US-10-237-551-193	Sequence 193, App
36	66.6	6.1	154746	10	US-09-827-688-8	Sequence 8, Appli
37	66.6	6.1	154746	10	US-09-827-688-8	Sequence 8, Appli
38	66.2	6.1	1614	9	US-09-976-740-45	Sequence 45, Appli
39	66.2	6.1	1614	13	US-10-023-529-45	Sequence 45, Appli
40	66.2	6.1	1614	13	US-10-023-523-45	Sequence 45, Appli
41	66.2	6.1	1614	17	US-10-616-187-45	Sequence 45, Appli
42	66.2	6.1	1614	17	US-10-671-242-45	Sequence 45, Appli
43	66.2	6.1	12425	9	US-09-976-740-50	Sequence 50, Appli
44	66.2	6.1	12425	13	US-10-023-529-50	Sequence 50, Appli
45	66.2	6.1	12425	13	US-10-023-523-50	Sequence 50, Appli

ALIGNMENTS

RESULT 1

US-10-672-399-5
; Sequence 5, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-5

Query Match 100.0%; Score 1083; DB 18; Length 1083;
Best Local Similarity 100.0%; Pred. No. 8.6e-257;
Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCCAGCCCGCGCTGCTCGGGCTGGCC	60
Db	1	ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCCAGCCCGCGCTGCTCGGGCTGGCC	60
Qy	61	CGGAGGGTGTGACCATCGCCCTTCGCGCTGCTCATCTGGGCTCATGACCTGGGCTTAC	120
Db	61	CGGAGGGTGTGACCATCGCCCTTCGCGCTGCTCATCTGGGCTCATGACCTGGGCTTAC	120
Qy	121	GCCGCCCGGGTGCCTGGCTCCGATCGCTACGGCTTCCTGGCTTCGGCTCTTACGGG	180

Db 121 GCCCGCGGGGTGCGGCTGGCTCGATCGCTACGGCCCTCCCTGGCGCTTCCGGCTCTACGGG 180
Qy 181 GCCTTCCTTTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTCGGGTACCTTGGAGACACGGCGG 240
Db 181 GCCTTCCTTTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTCGGGTACCTTGGAGACACGGCGG 240
Qy 241 GTGGCGGCGGCGCGCGGGCGCGCTGTGATGACGACACCGCGCGCAGTGTGGCGCTGACC 300
Db 241 GTGGCGGCGGCGCGCGGGCGCGCTGTGATGACGACACCGCGCGCAGTGTGGCGCTGACC 300
Qy 301 ATCTCCGCTTACCAGGAGACCCCGCTACCTGGCCAGTGCCTGGCGTCCGCCCGCGGCC 360
Db 301 ATCTCCGCTTACCAGGAGACCCCGCTACCTGGCCAGTGCCTGGCGTCCGCCCGCGGCC 360
Qy 361 CTGCTGTACCCGCGCGCGCTGGCGTCCCTCATGTGCTGATGAGTGGCAACCGCGCGGAG 420
Db 361 CTGCTGTACCCGCGCGCGCTGGCGTCCCTCATGTGCTGATGAGTGGCAACCGCGCGGAG 420
Qy 421 GACCTCTACATGCTGCAATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGCTAC 480
Db 421 GACCTCTACATGCTGCAATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGCTAC 480
Qy 481 GTGTGGACGGCAACTACCAACAGCCCTTGGGAACCCCGCGCGCGCGCGCTGGCGGCC 540
Db 481 GTGTGGACGGCAACTACCAACAGCCCTTGGGAACCCCGCGCGCGCGCGCTGGCGGCC 540
Qy 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCCTCGGCGCGCTTGGAGTGGAGCGCTGGTG 600
Db 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCCTCGGCGCGCTTGGAGTGGAGCGCTGGTG 600
Qy 601 AGGACTCGCAGGTGCGTGTGGCGCAGCGCTGGGGCGCGCAAGCGGAGTCTATGTAC 660
Db 601 AGGACTCGCAGGTGCGTGTGGCGCAGCGCTGGGGCGCGCAAGCGGAGTCTATGTAC 660
Qy 720 ACAGCCTTCAAGGCGCTCGAGATTCGGTGAGCTACGTGCAAGTCTGTGACTCGGACACA 720
Db 661 ACAGCCTTCAAGGCGCTCGAGATTCGGTGAGCTACGTGCAAGTCTGTGACTCGGACACA 720
Qy 721 AGGTGGACCCCATGGCACTGCTGGAGCTCTGTGGGGTACTGGACGAGGACCCCGGGTA 780
Db 721 AGGTGGACCCCATGGCACTGCTGGAGCTCTGTGGGGTACTGGACGAGGACCCCGGGTA 780
Qy 781 GGGCGTGTGTGGGGAGCTGGGATCTTTAAACCTCTGGACTCTGGGTGAGTTCCTA 840
Db 781 GGGCGTGTGTGGGGAGCTGGGATCTTTAAACCTCTGGACTCTGGGTGAGTTCCTA 840
Qy 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTCAAGACTACTTCCAC 900
Db 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTCAAGACTACTTCCAC 900
Qy 901 TGTGTATCTGCAATCAGCGTCTCTAGAACTCTGCCAGGCCCCAGGGACACGCGATG 960
Db 901 TGTGTATCTGCAATCAGCGTCTCTAGAACTCTGCCAGGCCCCAGGGACACGCGATG 960
Qy 961 ATGCCCTCATTCCTCGCCCCGTGCGAGTACACCTCCAGGTCCCGCTACTCAGAGAC 1020
Db 961 ATGCCCTCATTCCTCGCCCCGTGCGAGTACACCTCCAGGTCCCGCTACTCAGAGAC 1020
Qy 1021 GCCCTCGTCTTCTCGGGGTGGCTGAGCCAGCAGACGCTGTGCTCCAAAGTCTACTTCCG 1080
Db 1021 GCCCTCGTCTTCTCGGGGTGGCTGAGCCAGCAGACGCTGTGCTCCAAAGTCTACTTCCG 1080
Qy 1081 TGA 1083
Db 1081 TGA 1083

RESULT 2
US-10-672-399-3
; Sequence 3, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta

; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-3

Query Match 88.0%; Score 953.4; DB 18; Length 1071;
Best Local Similarity 94.5%; Pred. No. 6.3e-225;
Matches 1023; Conservative 0; Mismatches 1; Indels 59; Gaps 1;

Qy 1 ATGAGACAGCAGCAGCGCGCCCAAGCCACATCCCTGACGCCGCGCTGCTCGGCGCTCGCC 60
Db 1 ATGAGACAGCAGCAGCAGCGCGCCCAAGCCACATCCCTGACGCCGCGCTGCTCGGCGCTCGCC 60
Qy 61 CGGAGGGTGTGACCAATCGGCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 61 CGGAGGGTGTGACCAATCGGCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Qy 121 GC CGCGCGGGTGGCGCTGGCTTCGATCGCTACGGCTCTCTGGCTTCGGGCTCTACGGG 180
Db 121 GC CGCGCGGGTGGCGCTGGCTTCGATCGCTACGGCTCTCTGGCTTCGGGCTCTACGGG 180
Qy 181 GCCTTCCTTTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTCGCTACCTGGAGACACCGCGG 240
Db 181 GCCTTCCTTTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTCGCTACCTGGAGACACCGCGG 240
Qy 241 GTGGCGCGCGCGCGCGGGCGCTGGATGCAACACCGCGCGCAGTGTGGGCTCGCCCGCGCC 300
Db 241 GTGGCGCGCGCGCGCGGGCGCTGGATGCAACACCGCGCGCAGTGTGGGCTCGCCCGCGCC 300
Qy 301 ATCTCGCCTTAC CAGAGGACCCCGCGGTACTCGCGCGCAGTGTGGGCTCGCCCGCGCC 360
Db 301 ATCTCGCCTTAC CAGAGGACCCCGCGGTACTCGCGCGCAGTGTGGGCTCGCCCGCGCC 360
Qy 361 CTGCTGTACCCGCGCGCGCTGGCGTCCCTCATGTGCTGATGAGTGGCAACCGCGCGGAG 420
Db 361 CTGCTGTACCCGCGCGCGCTGGCGTCCCTCATGTGCTGATGAGTGGCAACCGCGCGGAG 420
Qy 421 GACCTCTACATGCTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGCTAC 480
Db 421 GACCTCTACATGCTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGCTAC 480
Qy 481 GTGTGGACCGGCAACTACCAACAGCCCTTGGGAACCCCGCGCGCGCGCTGGGCGGCC 540
Db 481 GTGTGGACCGGCAACTACCAACAGCCCTTGGGAACCCCGCGCGCGCGCTGGGCGGCC 540
Qy 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCCTGGGGCGCGCAAGCGGAGTCTATGTAC 600
Db 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCCTGGGGCGCGCAAGCGGAGTCTATGTAC 600
Qy 601 AGGACTCGCAGGTGCGTGTGGCGCAGCGCTGGGGCGCGCAAGCGGAGTCTATGTAC 660
Db 601 AGGACTCGCAGGTGCGTGTGGCGCAGCGCTGGGGCGCGCAAGCGGAGTCTATGTAC 660
Qy 661 ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTCTAGTGTGAGTCTGTGAGTCTCGGACACA 720
Db 661 ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTCTAGTGTGAGTCTGTGAGTCTCGGACACA 720
Qy 721 AGGTGGACCCCATGGCACTGCTGGGATCTTCGCGGGTACTGGACGAGGACCCCGGGTA 780
Db 721 AGGTGGACCCCATGGCACTGCTGGGATCTTCGCGGGTACTGGACGAGGACCCCGGGTA 780
Qy 781 GGGCGTGTGTGGGGAGCTGGGATCTTTAAACCTCTGGACTCTGGGTGAGTTCCTA 840
Db 781 GGGCGTGTGTGGGGAGCTGGGATCTTTAAACCTCTGGACTCTGGGTGAGTTCCTA 840

QY 841 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGACAGCTACTTCCAC 900
Db 841 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGACAGCTACTTCCAC 900
QY 901 TGTGTATCTGATCAGCGGCTCTCTAGATCCTGCCCCAGGCCCCAGGAGACCGGATG 960
Db 901 TGTGTATCTGATCAGCGGCTCTCT----- 926
QY 961 ATGCCCTCATCTCTCGCCCCCGTGCAGAGGTACACCTCCAGGTCCCGCTGCTACTCAGAGAC 1020
Db 927 -----AGGTACACTCCAGGTCCCGCTGCTACTCAGAGAC 961
QY 1021 GCCCTGCTCTCTGCGGTGCTGAGCCAGCAGACAGCTGGTCCAAAGTCTACTTCCG 1080
Db 962 GCCCTGCTCTCTGCGGTGCTGAGCCAGCAGACAGCTGGTCCAAAGTCTACTTCCG 1021
QY 1081 TGA 1083
Db 1022 TGA 1024

RESULT 3

US-10-672-399-1

; Sequence 1, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-1

Query Match 86.1%; Score 933; DB 18; Length 1737;
Best Local Similarity 90.6%; Pred. No. 6.2e-220;
Matches 1048; Conservative 0; Mismatches 35; Indels 74; Gaps 2;

QY 1 ATGAGACAGCAGGACGCGCCAGCCACTCTGCGAGCCGCGCTGCTCGGGCTGGCC 60
Db 1 ATGAGACAGCAGGACGCGCCAGCCACTCTGCGAGCCGCGCTGCTCGGGCTGGCC 60
QY 61 CGGAGGGTGTGACCATCGCCCTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db 61 CGGAGGGTGTGACCATCGCCCTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
QY 121 GCGCGCGGGTGCCTGCGCTCCGATCGCTACGCGCTCTGGGCTTCCGGCTTACCGG 180
Db 121 GCGCGCGGGTGCCTGCGCTCCGATCGCTACGCGCTCTGGGCTTCCGGCTTACCGG 180
QY 181 GCCTTCTTTACGCGCACCTGTGGCGCAGAGCTTTCGGCTACTGAGACACCGCGG 240
Db 181 GCCTTCTTTACGCGCACCTGTGGCGCAGAGCTTTCGGCTACTGAGACACCGCGG 240
QY 241 GTGGCGGCGGCGCGCGGGCGCTGTGATGCGACACCGCGCGAGTGTGGCGTGACC 300
Db 241 GTGGCGGCGGCGCGCGGGCGCTGTGATGCGACACCGCGCGAGTGTGGCGTGACC 300
QY 301 ATCTCCGCTTACAGGAGACCCCGCTAGCTGCGCCAGTGCCTTGGCGTCCGCGCGGCC 360
Db 301 ATCTCCGCTTACAGGAGACCCCGCTAGCTGCGCCAGTGCCTTGGCGTCCGCGCGGCC 360
QY 361 CTGCTGTACCCGCGCGCGGCTGCGCTCTCATGTGTGGATGGCAACCGCGCGGAG 420
Db 361 CTGCTGTACCCGCGCGCGGCTGCGCTCTCATGTGTGGATGGCAACCGCGCGGAG 420

QY 421 GACCTCTACATGGTGCACATGTTCCGCGAGGTCTTCTGCTGACGAGGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGGTGCACATGTTCCGCGAGGTCTTCTGCTGACGAGGACCCCGCCACGTAC 480
QY 481 GTGTGGGACGGCAATACACAGCCCTGGGAAACCCGCGGCGCGCGGCGCGGTGGGCGCC 540
Db 481 GTGTGGGACGGCAATACACAGCCCTGGGAAACCCGCGGCGCGCGGCGCGGTGGGCGCC 540
QY 541 GGAGCTATCGGAGGTGGAGCGGAGGATCCTGGGCGGCTGGCAGTGGAGCGCGCTGGTG 600
Db 541 GGAGCTATCGGAGGTGGAGCGGAGGATCCTGGGCGGCTGGCAGTGGAGCGCGCTGGTG 600
QY 601 AGGACTCCAGGTGCGTGTGCGCGCAGCGCTGGGCGGCAAGCGCGAGGTCAATGTAC 660
Db 601 AGGACTCCAGGTGCGTGTGCGCGCAGCGCTGGGCGGCAAGCGCGAGGTCAATGTAC 660
QY 661 ACAGCCTTCAAGGGCGCTCGGAGATTCGGTGGACTACGTCGAGGTCTGTGACTCGGACACA 720
Db 661 ACAGCCTTCAAGGGCGCTCGGAGATTCGGTGGACTACGTCGAGGTCTGTGACTCGGACACA 720
QY 721 AGGTTGGACCCCATGGCACTGCTGAGCTCGTGGGGTACTGGACGAGGACCCCGGGTA 780
Db 721 AGGTTGGACCCCATGGCACTGCTGAGCTCGTGGGGTACTGGACGAGGACCCCGGGTA 780
QY 781 GGGGCTGTTGTTGGGGAGCTGCGGATCCTTAACCTCTGGACTCTCTGGTCTGAGCTTCCCTA 840
Db 781 GGGGCTGTTGTTGGGGAGCTGCGGATCCTTAACCTCTGGACTCTCTGGTCTGAGCTTCCCTA 840
QY 841 AGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGACAGCTACTTCCAC 900
Db 841 AGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGACAGCTACTTCCAC 900
QY 901 TGTGTATCTGCTGATCAGCGGCTCTCTAG----- 928
Db 901 TGTGTATCTGCTGATCAGCGGCTCTCTAG----- 928
QY 929 -----AATCTGCCCCAGGCCCCAGGGAGCAGCGATGATGCC 965
Db 929 -----AATCTGCCCCAGGGTCTCTGGTACCACCTGTACTTTTGGGATGACCGG 1020
QY 966 CTCAATCTCTGCCCC-----CGTGCAGGTACACCTCAGGTCGCGC 1006
Db 966 CTCAATCTCTGCCCC-----CGTGCAGGTACACCTCAGGTCGCGC 1006
QY 1021 CACCTCAACCAACCGCATGCTCAGCATGGGTTATGTACCAAGTACACCTCCAGGTCCCGC 1080
Db 1021 CACCTCAACCAACCGCATGCTCAGCATGGGTTATGTACCAAGTACACCTCCAGGTCCCGC 1080
QY 1007 TGCTACTCAGAGACGCGCTCTCTCTGCGGTGGCTGAGCCAGCAGCAGCTGGTCC 1066
Db 1007 TGCTACTCAGAGACGCGCTCTCTCTGCGGTGGCTGAGCCAGCAGCAGCTGGTCC 1066
QY 1081 AAGTCGTACTTCCGTA 1083
Db 1081 AAGTCGTACTTCCGTA 1157

RESULT 4

US-10-042-523-1

; Sequence 1, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/10/042,523

APPLICATION NUMBER: US/10/042,523
FILING DATE: 19-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/635,552A
FILING DATE: 22-APR-1996
ATTORNEY/AGENT INFORMATION:

NAME: Brook, David E.
REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: LK95-07
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2116 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
LOCATION: 36..1769

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-042-523-1

Query Match 86.1%; Score 933; DB 13; Length 2116;
Best Local Similarity 90.6%; Pred. No. 6.1e-220;
Matches 1048; Conservative 0; Mismatches 35; Indels 74; Gaps 2;

QY 1 ATGAGACGAGGAGCGGCCCAAGCCCACTCTGTGAGCGCGCGCTGCTCGGCGCTGCGC 60
DB 36 ATGAGACGAGGAGCGGCCCAAGCCCACTCTGTGAGCGCGCGCTGCTCGGCGCTGCGC 95
QY 61 CGGAGGCTGTGACCATCGCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 96 CGGAGGCTGTGACCATCGCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 155
QY 121 GCGCGCGCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGG 180
DB 156 GCGCGCGCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGG 215
QY 181 GCCTTCCTTTAGCGGCACTGCTGCGCGAGAGCTCTTTCGGGTACCTGAGAGCAACCGCGCG 240
DB 216 GCCTTCCTTTAGCGGCACTGCTGCGCGAGAGCTCTTTCGGGTACCTGAGAGCAACCGCGCG 275
QY 241 GTGGCGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 276 GTGGCGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 335
QY 301 ATCTCCGCTTACAGAGAGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 336 ATCTCCGCTTACAGAGAGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 395
QY 361 CTGCTGTATCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 396 CTGCTGTATCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 455
QY 421 GACCTCTACATGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 456 GACCTCTACATGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 515
QY 481 GTGTGGGACGGCAACTACACAGAGCCCTGGGAAACCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 516 GTGTGGGACGGCAACTACACAGAGCCCTGGGAAACCGCGCGCGCGCGCGCGCGCGCGCGCG 575
QY 541 GGAGCCTATCGGAGGCTGAGAGCGGAGGATCTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 576 GGAGCCTATCGGAGGCTGAGAGCGGAGGATCTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 635

QY 601 AGGACTCGAGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 636 AGGACTCGAGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 695
QY 661 ACAGCCTTCAAGCGCTCGGAGATTTCGGTGGATCTACGTCAGAGTCTGTGATCTCGGACACA 720
DB 696 ACAGCCTTCAAGCGCTCGGAGATTTCGGTGGATCTACGTCAGAGTCTGTGATCTCGGACACA 755
QY 721 AGGTTGGACCCCATGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 756 AGGTTGGACCCCATGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
QY 781 GGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 816 GGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
QY 841 AGCAGCCTGCGATCTGCGGTAGCCTTCAATGTGAGCGGGCTTGTGAGAGTACTTTCAC 900
DB 876 AGCAGCCTGCGATCTGCGGTAGCCTTCAATGTGAGCGGGCTTGTGAGAGTACTTTCAC 935
QY 901 TGTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
DB 936 TGTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 995
QY 929 -----AATCTGCTCCAGGCCCCAGGGAGCAGCAGCAGATGATGCC 965
DB 996 CTTGAGGCTGGTACAAACAGAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055
QY 966 CTCATTCCTCGCCCC-----CGTCAAGGTACACTTCCAGGTCCCGC 1006
DB 1056 CACCTCAACCAACCGCATGCTCAGCATGGGTATGCTTACCAAGTACACTTCCAGGTCCCGC 1115
QY 1007 TGCTACTCAGAGAGCGCCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1066
DB 1116 TGCTACTCAGAGAGCGCCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1175
QY 1067 AAGTCGTACTTCCGTGA 1083
DB 1176 AAGTCGTACTTCCGTGA 1192

RESULT 5

US-10-672-399-7
; Sequence 7, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-7

Query Match 85.7%; Score 928; DB 18; Length 1065;
Best Local Similarity 100.0%; Pred. No. 1.1e-218;
Matches 928; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACGAGGAGCGGCCCAAGCCCACTCTGCTGAGCGCGCGCTGCTGCTGCTGCTGCTGCTG 60
DB 1 ATGAGACGAGGAGCGGCCCAAGCCCACTCTGCTGAGCGCGCGCTGCTGCTGCTGCTGCTGCTG 60
QY 61 CGGAGGCTGTGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 CGGAGGCTGTGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120

```
QY 121 GCCGCGGGGTCGCTGCGCTCCGATCCGATCCGCTCCTGGCTTGGCTCTACGGG 180
Db 121 GCCGCGGGGTCGCTGCGCTCCGATCCGATCCGCTCCTGGCTTGGCTCTACGGG 180
QY 181 GCCTTCTTTTACAGCGACCTGTGGCGAGAGCTTTCGCTACTGTGAGCACCGGG 240
Db 181 GCCTTCTTTTACAGCGACCTGTGGCGAGAGCTTTCGCTACTGTGAGCACCGGG 240
QY 241 GTGGCGGGCGCGCGGGGCGCTGGATGAGCACCACCGCGCAGTGTGGCGCTGACC 300
Db 241 GTGGCGGGCGCGCGGGGCGCTGGATGAGCACCACCGCGCAGTGTGGCGCTGACC 300
QY 301 ATCTCGGCTACAGAGAGACCCCGGTACTGCGCGAGTGTGCGCTGCGCGCGCC 360
Db 301 ATCTCGGCTACAGAGAGACCCCGGTACTGCGCGAGTGTGCGCTGCGCGCGCC 360
QY 361 CTGCTGTATACCCGCGCGGCTGCGCTCCTCATGGTGGTGGTAAACCGCGCGAG 420
Db 361 CTGCTGTATACCCGCGCGGCTGCGCTCCTCATGGTGGTGGTAAACCGCGCGAG 420
QY 421 GACCTCTACATGCTGACATGTTCCGCGAGGTCTTCGCTGACGAGAGACCCCGCAGTAC 480
Db 421 GACCTCTACATGCTGACATGTTCCGCGAGGTCTTCGCTGACGAGAGACCCCGCAGTAC 480
QY 481 GTGGGAGCGCAACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGTGGCGCC 540
Db 481 GTGGGAGCGCAACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGTGGCGCC 540
QY 541 GGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGGT 600
Db 541 GGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGGT 600
QY 601 AGGACTCGCAGTGTGCTGCGCGAGCGCTGGGCGGCAAGCGCGAGTCACTGATAC 660
Db 601 AGGACTCGCAGTGTGCTGCGCGAGCGCTGGGCGGCAAGCGCGAGTCACTGATAC 660
QY 661 ACAGGCTTCAAGCGCGTGGAGATTCGCTGAGTACGTCAGGTCGTGACCTCGACACA 720
Db 661 ACAGGCTTCAAGCGCGTGGAGATTCGCTGAGTACGTCAGGTCGTGACCTCGACACA 720
QY 721 AGGTTGGACCCATGCACTGCTGAGCTCGTGGGCTGCTGAGAGACCCCGGGTA 780
Db 721 AGGTTGGACCCATGCACTGCTGAGCTCGTGGGCTGCTGAGAGACCCCGGGTA 780
QY 781 GGGGCTGTGTGGGAGATCGGATCCCTTAACCTCTGGAATCTGCTGCTGAGCTTCTTA 840
Db 781 GGGGCTGTGTGGGAGATCGGATCCCTTAACCTCTGGAATCTGCTGCTGAGCTTCTTA 840
QY 841 AGCAGCTGCGATACCTGGTAGCCTTCAATGTGGAGCGGCTTGTGACAGCTACTTCCAC 900
Db 841 AGCAGCTGCGATACCTGGTAGCCTTCAATGTGGAGCGGCTTGTGACAGCTACTTCCAC 900
QY 901 TGTGTATCTGCATCAGCGGCTCTCTAG 928
Db 901 TGTGTATCTGCATCAGCGGCTCTCTAG 928
```

RESULT 6

US-09-902-939-1

```
; Sequence 1, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
```

RESULT 7

```
; ORGANISM: mouse
US-09-902-939-1
```

Query Match 62.6%; Score 677.8; DB 10; Length 1752;

Best Local Similarity 84.5%; Pred. No. 3.4e-157; Indels 18; Gaps 2;
Matches 792; Conservative 0; Mismatches 127;

```
QY 10 CAGGACGGCCCAAGCCCACTCTCTCAGCCCGCCGCTGCTCGGCTTGGCCCGGAGGTTG 69
Db 7 CAGGACATGCAAAAGCCCTCAGAGCAGCGGTTGCTCTGCTGCTGCTGCTGCTGCTGCTG 66
QY 70 CTGACCATCGCTTTCGCGCTGCTCTATCTTGGGCTCATGACCTGGGCTTACCGCCGCG 129
Db 67 CTCACGATCATCTTTGCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTACCGCCGAG 126
QY 130 GTGCGCTGCGCTTCCGATCGCTTACCGCTTCTGGGCTTCTGGGCTTCTACGGGCTTCT 189
Db 127 GTTCTCTGCTTCTGATCGCTATGATCTCTCTGGGCTTCTGGGCTTCTATGGGCTTCT 186
QY 190 TCAGCGCACCTGTGGCGCAGAGCTTCTCGCTACCTGGAGCACCGCGGCTGGCGG 249
Db 187 AGCGCACACCTAGTGGCACAGAGCTTCTTCTTACCTGGAGCACCGAAGGCTGGCAGCG 246
QY 250 GGGGCGCG-----GGGCGCGCTGGATGACGACCCCGCGGCTGCTGGCG 294
Db 247 GCTGCGCGCGCTCTCTTGGCGAAGGGGCTTGGATGCGGCGCTGCTGCTGCTGCTGCTG 306
QY 295 CTGACCATCTCGCGCTTACAGAGAGACCCCGCTACCTGCGGCTGCTGCTGCTGCTGCTG 354
Db 307 CTCACCATCTCAGCTTACAGAGATCTCCGCTTACCTGCGGCTGCTGCTGCTGCTGCTG 366
QY 355 CGCGCTTCTGCTTACCGCGCGCGCTGCGGCTCTCATGCTGGTGGATGGCAACCGC 414
Db 367 CGCGCTTCTGCTTACCGCGCACACGAGTTACGCTGCTCATGCTGGTGGATGGCAACCGC 426
QY 415 GCGAGGACCTCTACATGCTGCAATGTTCCGCGAGGCTCTTCTGCTGACGAGACCCCGCC 474
Db 427 GCTGAGGATCTGATACATGCTGACATGTTCCGAGAGTCTTCCGCGATGAGGACCCCGCC 486
QY 475 ACCTAGCTGTGGGACCGCAATACCAACAGCCCTGGGAAACCGCGCGCG-- --CGGCGCG 531
Db 487 ACTATGCTGGGATGGCACTACCATCAGCCCTGGGAAACCGAGAGCTACGCGCGCT 546
QY 532 GTGGCGCGCGGACCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAG 591
Db 547 GTCGCTGAAGGTGCTTACCGGAGGTGGAGCGGAGGACCCCGGCGGTTGGCGGTGGAG 606
QY 592 GCGCTGTGTGAGGACTCGCAGGTGCTGCTGGCGAGCGCTGGGCGGCGGAGCGCGAG 651
Db 607 GCGCTGTGTGAGAACACACGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
QY 652 GTCATGTACACAGCTTCAAGCGCTCGAGATTCGGTGGACTACGTCGAGTCTGTGAC 711
Db 667 GTCATGTACACAGCTTTCAGGCACTGGCGGACTCCGCTGGAGTCTGCTGAGTCTGTGAC 726
QY 712 TCGGACACAAGGTTGGACCCCATGCTGCTGGAGCTGCTGGGCTTACTGGACGAGGAC 771
Db 727 TCAGACACAAGACTAGACCCCATGGCACTGCTGGAGCTTGTGGAGTGTGGATGAAGAC 786
QY 772 CCGCGGTAGGGCTGTTGGTGGGAGCTGCGGATCTTAAACCTCTCTGAGCTCTGGGTC 831
Db 787 CCGCGGTAGGGCTGTTGGAGGGGATGTGAGGATCTTAAACCTCTCTGAGCTCTGGGTC 846
QY 832 AGCTTCTTAAGCAGCTCGGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGGAC 891
Db 847 AGCTTCTTGAAGAGCTTCTCGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGGAC 906
QY 892 TACTTCCACTGTGTATCTCTGCATCAGCGGCTCTCTAG 928
Db 907 TACTTCCACTGTGTGTCTCTGCATCAGTGGTCTCTCTAG 943
```

US-10-363-345A-13975
; Sequence 13975, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 13975
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 13975
US-10-363-345A-13975

Query Match 39.0%; Score 422; DB 18; Length 662;
Best Local Similarity 77.3%; Pred. No. 2.8e-94;
Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY 34 GGAGCCCGCGTGTCTCCGGCTGCGCCCGGAGGGTGTGACATCGCCTTCGCCCTGCTC 93
Db 1 GTAGTTTCTCGTTGTTCCGTTGCTCGGAGGGTGTGATATCGTTTCGTTTGTGTT 60
QY 94 ATCTGGGCTCATGACCTGGGCTACCGCGCGGGGTGCGCTGGCCTCGATCGCTAC 153
Db 61 ATTTTGGGTTTATGATTTGGGTTTACGTCGCGGGGTGCTGTTGGTTTCGATCGTTAC 120
QY 154 GGCCTCTCTGGCTTCGGCTCTACCGGGCTTCCTTTTCAGGCACTCGTGGCGCAGAGC 213
Db 121 GGTTTTGTGTTTCGGTTTACGGGTTCCTTTTATAGCGTATTTGGTGGCGTAGAGT 180
QY 214 CTCTTCGGCTACCTGGAGACCGCGGGGTGCGCGGGCGCGGGCGCGCTGATGCA 273
Db 181 TTTTTCGGCTATTTGGAGTATCGCGGGTGGCGCGGGCGCGGGGTCTGTTGATGTA 240
QY 274 GCCACCGCGCAGTGTGCGCTGACCATCTCCGCTACAGGAGGACCCCGCGTACCTG 333
Db 241 GTTATCGCGTAGTGTGCGCTGATTTATTTTCGTTTATAGAGGATTTCCGCTATTG 300
QY 334 CGCCAGTGCCTGGCTCCGCGCGCTGTGTACCGCGCGCGCGCTGCGCTCCTC 393
Db 301 CGTTAGTGTTCGCGTTCTGTTTCGCGTTTGTGTTATTCGCGCGCGGTTGGCGTTT 360
QY 394 ATGTTGTGATGGCAACCGCGCGCAGACCTCTACATGTCGATGTCGATGTCGCGAGTC 453
Db 361 ATGTTGTGATGGTAAATCGCGCTCGAGGATTTTATATGTCGATATGTTTCGCGAGTT 420
QY 454 TTGCTGACGAGGACCCCGCAGTCTGTTGGAGCGCACTACACCGCTCGGAA 513
Db 421 TTTGTTGACGAGGATTTCTGTTACGTCGTTGGAGCGGTAAATTTATTTAGTTTGGAA 480
QY 514 CCGCGCGCGCGCGCGCGTGGCGCGCAGAGCTATCGGAGGTGGAGCGCGAGGATCCT 573
Db 481 TTCGCGCGCGCGCGCGTGGCGTTCGAGTTTATCGGAGGTGGAGCGCGAGGATTTT 540
QY 574 GGGCGGTGGCAGTGGAGCGCTGTGAGGACTCGCAGGTGCTGTCGTCGCGCAGCGC 633
Db 541 GGGCGGTGGTAGTGGAGCGCTGTGAGGATTCGTAGGTGCTGTCGTCGCGTAGCGT 600
QY 634 TGGGCGCGCAGCGAGGTATGATACAGCTTCAGAGGCTCGGAGATTCGTTGGAC 693
Db 601 TGGGCGGTGAGCGCAGGTTATGATATAGTTTATAGGGCTTCGGAGATTCGTTGGAT 660
QY 694 TA 695
Db 661 TA 662

RESULT 8
US-10-363-345A-13976/c
; Sequence 13976, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 13976
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 13976
US-10-363-345A-13976

Query Match 39.0%; Score 422; DB 18; Length 662;
Best Local Similarity 77.3%; Pred. No. 2.8e-94;
Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY 34 GCAGCCCGCGCTGCTCCGGCTGCGCCCGGAGGGTGTGACATCGCCTTCGCCCTGCTC 93
Db 662 GTAGTTTGTGTTGTTTCGTTTGGTTTCGAGGGTGTGATATCGTTTCGTTTGTGTT 603
QY 94 ATCTTCGGCTCATGACCTGGGCTACCGCGCGGGGTGCGCTGGCCTCGATCGCTAC 153
Db 602 ATTTTGGGTTTATGATTTGGGTTTACGTCGCGGGGTGCTGTTGGTTTCGATCGTTAC 543
QY 154 GGCCTCTCTGGCTTCGGCTCTACGGGCTTCCTTTTCAGGCACTCGTGGCGCAGAGC 213
Db 542 GGTTTTGTGTTTCGGTTTTCGCGGTTCCTTTTATAGCGTATTTGGTGGCGTAGAGT 483
QY 214 CTCTTCGGCTACCTGGAGACCGCGCGGTGCGCGCGCGCGCGCGCTCGATGCA 273
Db 482 TTTTTCGGTATTTGGAGTATCGCGGTGCGCGCGCGCGCGCGGTCTGTTGATGTA 423
QY 274 GCCACCGCGCAGTGTGCGCTGACCATCTCCGCTACAGGAGGACCCCGCGTACCTG 333
Db 422 GTTATCGCGTAGTGTGCGCTTCGATTTATTTTCGTTTATAGGAGGATTTCCGCTATTG 363
QY 334 CGCCAGTGCCTGGCGTCCGCGCGCTGTCGTACCGCGCGCGCGCTGCGGCTCCTC 393
Db 362 CGTTAGTGTTCGCGTTTCGTTGTTGTTATTCGCGCGCGCGGTTGGCGTTT 303
QY 394 ATGTTGTGATGGCAACCGCGCGCAGACCTCTACATGTCGATGTCGCGAGTC 453
Db 302 ATGTTGTGATGGTAAATCGCGCTCGAGGATTTTATATGTCGATATGTTTCGCGAGTT 243
QY 454 TTGCTGACGAGGACCCCGCAGTCTGTTGGAGCGCACTACACCGCTCGGAA 513
Db 242 TTTGTTGACGAGGATTTCTGTTACGTCGTTGGAGCGGTAAATTTATTTAGTTTGGAA 183
QY 514 CCGCGCGCGCGCGCGTGGCGCGCAGAGCTATCGGAGGTGGAGCGCGAGGATCCT 573
Db 182 TTCGCGCGCGCGCGCGTGGCGTTCGAGTTTATCGGAGGTGGAGCGCGAGGATTTT 123
QY 574 GGGCGGTGGCAGTGGAGCGCTGTTGAGGACTCGCAGGTGCTGTCGTCGCGCAGCGC 633
Db 122 GGGCGGTGGTAGTGGAGCGCTGTTGAGGATTCGTAGGTGCTGTCGTCGCGTAGCGT 63
QY 634 TGGGCGCGCAGCGAGGTATGATACAGCTTCAGAGGCTCGGAGATTCGTTGGAC 693
Db 62 TGGGCGGTGAGCGCAGGTTATGATATAGTTTATAGGGCTTCGGAGATTCGTTGGAT 3

QY 694 TA 695
Db 2 TA 1

RESULT 9

US-10-363-345A-13973/c

; Sequence 13973, Application US/10363345A

; Publication No. US20040234960A1

; GENERAL INFORMATION:

; APPLICANT: Alexander Olek

; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Method for determining the degree of methylation of defined

; FILE REFERENCE: E01/1227

; CURRENT APPLICATION NUMBER: US/10/363,345A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 13973

; LENGTH: 662

; TYPE: DNA

; ORGANISM: Artificial Sequence

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 13973

US-10-363-345A-13973

Query Match 36.4%; Score 393.8; DB 18; Length 662;

Best Local Similarity 74.7%; Pred. No. 2.4e-87;

Matches 494; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 35 CAGCCGCGCTGCTCCGCGCTGCGCGGAGGTGCTGACCATCGCTTCGCGCTGCTCA 94

Db 661 CAACCTACCGCTACTCCGACCTAACCCGAAATACTACCATCGCTTCGCGCTACTCA 602

QY 95 TCTGGGCTCATGACTGGGCTACCGCGCGGGGTGCGGTGCGCTCCGATCGCTAG 154

Db 601 TCTAAACCTCACTAAACCTACCGCGCGGAAATACCGGTAACTCCGATCGCTAG 542

QY 155 GCTCTGCGCTTCGCGCTTACGGGCTTCTGCGGCTTCTTTCAGCGACCTGGTGGCGAGGC 214

Db 541 ACCTCTAACCTTCGACCTTCTAGAAACCTTCTTTCAGCGACCTTAATACGCAAAAC 482

QY 215 TCTTCGCTACCTGGAGCACCGCGCGGTGCGCGCGCGCGCGCTGGATGCGAG 274

Db 481 TCTTCGCTACCTAAACACCGAGATACGACGACGCGGAAACCGCTAAATACAA 422

QY 275 CCACCGCGCGAGTGTGGCGCTGACCATCTTCGCGCTTACAGGAGGACCCCGGTACCTGC 334

Db 421 CCACCGCGCGCAATATAAGCTTAACCATCTCGCGCTTACCAAAAAACCCCGGTACCTAC 362

QY 335 GCCAGTGCCTGCGCTCCGCGCGCGCTGCTGACCGCGCGCGCGCGCTGCGGTCTCA 394

Db 361 GCCAATACCTAACGTCGCGCGCGCTTACTATACCGCGCGCGCGGACTACGCGTCTCA 302

QY 395 TGTGTGTGATGGCAACCGCGCGAGGACTCTACATGCTGCACATGTTCCGCGAGTCT 454

Db 301 TAATAATAATAACACCGCGCGGAAACCTCTACATAATCGACATATTCGCGGAAATCT 242

QY 455 TCGCTGACGAGGACCCCGCAAGTGTGTGGGACGGCACTACCAACGCGCTGGGAG 514

Db 241 TCGCTAAACGAAACCCCGCGCAGTACGTATAAACGACAACTACCAACCAACCTAAAC 182

QY 515 CCGCGCGCGCGCGGTGGGCGCGGAGCTTATCGGAGGTGGGCGGAGGATCTCTG 574

Db 181 CCGGACGACGAGCGGATTAACCGCGGAAACCTTATCGAAAAATAAAACGAAAAATCTTA 122

QY 575 GCGCGTGCAGTGGAGGCGCTGGTGGAGCTCGCAGGTGCGGTGCGGTGGCGAGCGCT 634

Db 121 ACGACTAATCAATAAAACGCTAATAAAACCTTCGCAATACGTATACGTAACGCAACGCT 62

QY 635 GGGCGGCAAGCGCGAGGTCTATGTATACAGCCTTCAAGCGCGCTCGGAGATTGCGTGGACT 694

Db 61 AAACCGACAAACCGGAATCATATACAAACCTTCAAAACGCTCGAATAATCGATAACT 2
QY 695 A 695
Db 1 A 1

RESULT 10

US-10-363-345A-13974

; Sequence 13974, Application US/10363345A

; Publication No. US20040234960A1

; GENERAL INFORMATION:

; APPLICANT: Alexander Olek

; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Method for determining the degree of methylation of defined

; FILE REFERENCE: E01/1227

; CURRENT APPLICATION NUMBER: US/10/363,345A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 13974

; LENGTH: 662

; TYPE: DNA

; ORGANISM: Artificial Sequence

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 13974

US-10-363-345A-13974

Query Match 36.4%; Score 393.8; DB 18; Length 662;

Best Local Similarity 74.7%; Pred. No. 2.4e-87;

Matches 494; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 35 CAGCCGCGCTGCTCCGCGCTGCGCGGAGGTGCTGACCATCGCTTCGCGCTGCTCA 94

Db 2 CAACCTACCGCTACTCCGACCTAACCCGAAATACTACCATCGCTTCGCGCTACTCA 61

QY 95 TCTTCGCGCTCATGACTGGGCTTACCGCGCGGGGTGCGGTGCGCTCCGATCGCTAG 154

Db 62 TCTTAACCTCATAACTTAACCTACCGCGCGGAAATACCGTAACTTCGATCGCTAG 121

QY 155 GCTCTGCGCTTCGCGCTTACGCGGCTTCTTTCAGCGCACCTGTGGGCGAGGC 214

Db 122 ACCTCTTAACCTTTCGACCTTCTAGAAACCTTCTTTCAGCGACCTTAATACGCAAAAC 181

QY 215 TCTTCGCGTACTCGGAGCACCGCGGGTGGCGGGCGCGCGCGCGCTGGATGCGAG 274

Db 182 TCTTCGCGTACTTAAACACCGACGAAATAACGACGACGCGGAAACCGCTTAATACAA 241

QY 275 CCACCGCGCGAGTGTGGGCTGACCATCTTCGCGCTTACAGGAGGACCCCGGTACCTGC 334

Db 242 CCACCGCGCGCAATATAACGCTTAACCTTCGCGCTTACCAAAAAACCCCGGTACCTAC 301

QY 335 GCCAGTGCCTGCGCTCCGCGCGCTGCTGATCTGCTACCGCGCGCGGTGCGGTCTCA 394

Db 302 GCCAATACCTTAACGTCGCGCGCGCTTACTATACCGCGCGCGGACTACGCGTCTCA 361

QY 395 TGTGTGTGATGGCAACCGCGCGGAGGACTCTTCTACATGCTGCACATGTTCCGCGAGTCT 454

Db 362 TAATAATAATAACCAACCGCGCGGAAACCTCTACATAATCGACATATTCGCGGAAATCT 421

QY 455 TCGCTGACGAGGACCCCGCAAGTGTGTGGGACGGCACTACCAACGCGCTGGGAG 514

Db 422 TCGCTAAACGAAACCCCGCGCAGTACGTATAAACGACAACTTACCAACCAACCTAAAC 481

QY 515 CCGCGCGCGCGCGGTGGGCGCGGAGCTTATCGGAGGTGGGCGGAGGATCTCTG 574

Db 482 CCGGACGACGAGCGGATTAACCGCGGAAACCTTATCGAAAAATAAAACGAAAAATCTTA 541

QY 575 GCGCGTGCAGTGGAGGCGCTGGTGGAGACTTCGAGGTGCGGTGCGGTGGCGAGCGCT 634

Db 542 AACGACTAACATAAAAAACGCTAATAAAAACTCGCAATACGTATATACGTAAACGCAACGCT 601
Qy 635 GGGGGCGCAAGCGGAGGTCTGTACACAGCCTTCAAGCGCTCGGAGTTCGGTGGACT 694
Db 602 AAACGACAAACGGAATCATATACAAACCTTCAAAACGCTCGAAATTCGATAAACT 661
Qy 695 A 695
Db 662 A 662

RESULT 11

US-09-918-995-27210
; Sequence 27210, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27210

Query Match 31.3%; Score 339.2; DB 10; Length 490;
Best Local Similarity 91.7%; Pred. No. 6.5e-74;
Matches 353; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
Qy 544 GCCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAGGCGCTGGTGAGG 603
Db 2 GCAGGTAGGGAANNNNNNNAATTGAGACCTGAGGCACNGTCGGGGAANTCGTGGTGAGG 61
Qy 604 ACTCGAGGTGGTGTGGTGGCGCAGCGCTGGGGCGGCAAGCGGAGGTCTATGTACACA 663
Db 62 ACTCGAGGTGGTGTGGTGGCGCAGCGCTGGGGCGGCAAGCGGATGTATGTACACA 121
Qy 664 GCCTTCAAGGCGCTCGGAGATTTCGGTGACTACGTGACGTTGAGTCTGGACACAAAGG 723
Db 122 GCCTTCAAGGCGCTCGGAGATTTCGGTGACTACGTGACGTTGAGTCTGGACACAAAGG 181
Qy 724 TTGACACCCCATGGCACTGCTGGAGCTCGTGGGGTACTTGGACGAGGACCCCGGGTAGGG 783
Db 182 TTGACACCCCATGGCACTGCTGGAGCTCGTGGGGTACTTGGACGAGGACCCCGGGTAGG 241
Qy 784 GCTGTGTGGGGAAGCTGGGAGATCTTAAACCTCTGGAATCTCTGGGTGAGTCTTCAAGC 843
Db 242 GCTGTGTGGGGAATGTGGGATCTTAAACCTCTGGACTCTCGGGTCAAGTCTTCAAGC 301
Qy 844 AGCTGCGCATACTGGGTAGCCCTTCAATGTGGAGCGGCTTGTGACAGCTACTTCCACGT 903
Db 302 AGCTGCGCATACTGGGTAGCCCTTCAATGTGGAGCGGCTTGTGACAGCTACTTCCACGT 361
Qy 904 GTATCCTGCATCAGCGGTCCTCTAG 928
Db 362 GTATCCTGCATCAGCGGTCCTCTAG 386

RESULT 12

US-10-295-027-369
; Sequence 369, Application US/10295027
; Publication No. US20030232350A1

; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 369
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-369

Query Match 25.2%; Score 273; DB 17; Length 1662;
Best Local Similarity 61.1%; Pred. No. 1.1e-57;
Matches 532; Conservative 0; Mismatches 300; Indels 39; Gaps 4;
Qy 61 CGGAGGTGTGACCATCGCTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db 28 CGTGTGTGGGCACACAGCTGTGTTGCCCTGGCAGTGTGGGTGGCATCTGGCAGCCTAT 87
Qy 121 GCCCGCGGGGTGCGCTGGCCTCCGATCGCTACGCGCTCTCGGCTTCGGCTCTACGGG 180
Db 88 GTGACGGGCTACAGTTTCATCCACGGAAGACACTACCTGTCTTCGGCTGTACGGC 147
Qy 181 GCCTTCCTTTTCAGCGCACTGGTGGCGCAGAGCTCTTTCGGGTACTCTGGAGACCGGGCG 240
Db 148 GCCATCTCGGCTGTGACCTGCTCATTCAGAGCCTTTTTCGCTTCTCTGGAGACCGGGCG 207
Qy 241 GTGGCGGGCGGGCGGGGGCGGCTGGATGACGACACCGGCGC---GCAGTGTGGCGGTG 297
Db 208 ATGCGACGTGCGCGCCAGGGCCCTGAAGTGCCTCCCGCGGGCGGGCTCGGTGGCACTG 267
Qy 298 ACCATCTCCGCTACACAGGAGGACCCCGCTACCTCGGCCAGTGCCTGGCGCTCGCGCCGCG 357
Db 268 TGCATTGCGCATACAGGAGGACCTTACTTGTGGCAAGTGCCTGGCTCGGCTCGGCCAG 327
Qy 358 GCCTGTGTGTACCGCGCGGGCGGCTCGGGTCTCTCATGTGTGTGGATGGCAACCGCGCC 417
Db 328 GCATCTCTTCCCTGAC-----CTCAAGGTGTGTGTGTGTGGATGGCAACCGGCCAG 381

QY 418 GAGGACCTTACATGTGACATGTTTCCGCGAGGCTTTCGTGACGAGACCCCGCCACG 477
|||||
Db 382 GAGGAGCCCTACATGTGACATCTTCCACGAGTGTCTGGCGGACACCGAGCAGCGCGC 441
QY 478 TACGTGTGGAGCGGCAACTACACACAGCCCTGGGAACCCGCGCGCGCGCGGTGGGC 537
|||||
Db 442 TTCTTTGTGTGGCGCAGCAACTTCCA-----TGAGGCGAGCGGAGGTGAG 486
QY 538 GCCGAGCCCTATCGGAGGTGAGGCGGAGGATCTTGGCGGCTGGCACTGGAGGCGCTG 597
|||||
Db 487 ACGGAGCCAGCTGACGAGGCGCATGACCGT-----GTGCGGGATGTG 531
QY 598 GTGAGACTCGCAGGTGCTGTGCGTGGCGCAGCGCTGGGCGGCGCAAGCGCAGGTCTATG 657
|||||
Db 532 GTGCGGCGCAGCACCTTCTCGTGCATCATGCAAGTGGGAGGCAAGCGCGAGGTCTATG 591
QY 658 TACAGCAGCTTCAAGCGCTCGGAGATTGCGTGGACTAGTGCAGCTGTGACTCGGAC 717
|||||
Db 592 TACAGCGCTTCAAGCGCTCGGAGATTGCGTGGACTAGTGCAGCTGTGACTCGGAC 651
QY 718 ACAAGGTTGGACCCCATGCACTGCTGGAGCTCGTGGCGGTACTGGACGAGGACCCCGCG 777
|||||
Db 652 ACTGTGTGATCCAGCTCGACCATCGAGATGCTTCGAGTCTTGGAGGAGGATCCCCAA 711
QY 778 GTAGGGCTGTGGTGGGACGTGGGATCTTAAACCTCTCGACTCTCGGCTCGAGTTC 837
|||||
Db 712 GTAGGGGAGTGGGGAGATGTCCAGATCTCTCAAGTACGACTCATGGATTTCTCTTC 771
QY 838 CTAAGCAGCTCGATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 897
|||||
Db 772 CTGACAGCGCTCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
QY 898 CACTGTGTATCTGTGATCAGCGCTCTCTAG 928
|||||
Db 832 GCCTGTGTGCTGTATGATGAGGCGCTTGG 862

RESULT 13

US-10-188-832-136
; Sequence 136, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 136
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-188-832-136

Query Match

Best Local Similarity 25.2%; Score 273; DB 17; Length 1662;
Matches 532; Conservative 0; Mismatches 300; Indels 39; Gaps 4;

RESULT 14

US-10-309-560-9
; Sequence 9, Application US/10309560
; Publication No. US20030235893A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: KUMARI, KSHAMA
; TITLE OF INVENTION: HYALURONAN SYNTHASES AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 35541.082
; CURRENT APPLICATION NUMBER: US/10/309,560
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/336,105

QY 61 CGAGGGGTGTGACCATCGCCTTCCCTGTGCTATCTTGGGCTCATGACCTTGGGCTTAC 120
|||||
Db 28 CGTGTGTGGGACACAGCCTGTTTCCCTTGGCAGTGTGCTGGTGGCATCTTGGCAGCCTAT 87
QY 121 GCCGCGGGGTGCGCCTGCTCCGATCGCTAGGGCTCTTGGCTTGGCTTGGCTTGGCTTGG 180
|||||
Db 88 GTGACGGGTACCACTTATCCACACGGAAAGCACTACCTGTCTCTTGGCTTGGCTTGG 147
|||||
QY 181 GCTTCTCTTTCAGCGCAGCCTGTGTGGCGCAGAGCCTCTTTCGCTTACCTGTGAGCACC 240
|||||
Db 148 GCCATCTCTGGGCTGACCTGCTCATTCAGAGCCTTTTGGCTTCTTGGAGCACC 207
QY 241 GTGCGGCGGCGCGCGCGGCGCGCTGTGATGAGCAGCAACCGCGC-----GCAGTGTGG 297
|||||
Db 208 ATGCGACGCTGCGCGCGCAGGCGCTGAAGCTGCCCTCCCGCGCGGCGGTCTCGTGG 267
QY 298 ACCATCTCGGCTTACAGGAGGAGACCCCGGTACTCTGCGCAGTGTGCGCTGCGCTGCG 357
|||||
Db 268 TGCATTGCGCGGTACAGGAGGACCTGACTACTTTGGCGAAGTGTGCTGCGCTGCGGCC 327
QY 358 GCGCTGTGTATCCCGCGCGCGCGCTGGCGCTCTCATGCTGTGTGATGGCAACCGCGCTC 417
|||||
Db 328 CGCATCTCTTCTCCCTGAC-----CTCAAGTGTGTGATGGTGGTGGTGGTGGTGG 381
QY 418 GAGGACCTTACATGTGACATGTTCCGCGAGGTCTTTCGTGACGAGACCCCGCCACG 477
|||||
Db 382 GAGGAGCGCTACATGCTGACATCTTCCAGAGGTGCTGGGCGGCAACCGAGCAGCGCGC 441
QY 478 TACGTGTGGAGCGGCAACTACACACAGCCCTGGGAAACCCCGCGCGCGCGCGGTGGGC 537
|||||
Db 442 TTCTTTGTGTGGCGCAGCAACTTCCA-----TGAGGCGAGGCGGAGGTGAG 486
QY 538 GCGGAGCCTATTCGGGAGGTGGAGCGGAGGATCTTGGCGGCTGSCAGTGGAGGCGCTG 597
|||||
Db 487 ACGGAGCGCAGCTTCAGAGGCGCATGACCGT-----GTGCGGGATGTG 531
QY 598 GTGAGGACTCGCAGGTGCTGTGCGTGGCGCAGCGCTGGGCGGCGCAAGCGCAGGTCTATG 657
|||||
Db 532 GTGCGGCGCAGCACCTTCTCGTGCATCATGCAAGAGTGGGAGGCAAGCGCAGGTCTATG 591
QY 658 TACAGCAGCTTCAAGCGCTCGGAGATTGCGTGGACTAGTGCAGGTCTGTACTCGGAC 717
|||||
Db 592 TACAGCGCTTCAAGCGCTCGGCGATTCGGTGGACTACATCCAGGTGTGCGACTCTGAC 651
QY 718 ACAAGGTTGGACCCCATGCACTGCTGGAGCTCGTGGCGGTACTGGACGAGGACCCCGCG 777
|||||
Db 652 ACTGTGTGATCCAGCTCGACCATCGAGATGCTTCGAGTCTTGGAGGAGGATCCCCAA 711
QY 778 GTAGGGCTGTGTGGGAGCTGCGGATCTTAAACCTCTTGGACTCTCGGCTCGAGTTC 837
|||||
Db 712 GTAGGGGAGTGGGGAGATGTCCAGATCTCTCAAGTACGACTCATGGATTTCTCTTC 771
QY 838 CTAAGCAGCTCGATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 897
|||||
Db 772 CTGACAGCGCTCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
QY 898 CACTGTGTATCTGTGATCAGCGCTCTCTAG 928
|||||
Db 832 GCCTGTGTGCTGTATGATGAGGCGCTTGG 862

```
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Xenopus laevis
US-10-309-560-9

Query Match      23.1%; Score 250; DB 17; Length 1767;
Best Local Similarity 55.8%; Pred. No. 5e-52;
Matches 503; Conservative 0; Mismatches 390; Indels 9; Gaps 1;

QY 62 GAGAGGTGCTGACCATCGCTTCGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAG 121
DB 80 GGAGGAAATATTTATTTCTTTGGTGGTCTATTAGCTACCAATACAGCAGCTTAG 139
QY 122 CGCGCGGGGTGCGCTGGGCTCCGATCGCTACGGCTCTCGGCTTCGGCTCTACGGG 181
DB 140 TGGCAGAGTTCAGGTCTCAACATGAAGCATTTCTCTCTCCCTTGGGCTTTATGTC 199
QY 182 CTTCTCTTTCAGCGCACTGTGGCGCAGAGCTCTTCGCGTACCTGGAGCACCGCGGG 241
DB 200 TTGCAATGCTTCTCCACCTGATGATGACAGAGCTCTTTGGCTTCTCGGAGATACGAGG 259
QY 242 TGGCGCGCGCGCGCGGGGCGCTGGATGACGCCACCGCGGCTAGTGTGGCTGACCA 301
DB 260 TAAATAAGAGTGA-----GCTTCTCTGCGCTTTTAAAGAACAGCTGGCTCTGACCA 310
QY 302 TCTCCGCTTACCAGAGGACCCGCGGTACTCTCGCGCAGTGTGCGCTCCGCGCGCGCC 361
DB 311 TTGCTGGGTATCAGAGAACCTTGATCTTCTGTAAGTGTGTAATCTTGCAGATAG 370
QY 362 TGCTGTACCGCGCGCGGCTGCGGCTCTCATGGTGTGATGGCAACCGCGCGGAGG 421
DB 371 TGAATAACCCCAAGATAAATCAAGATCATTTTGGTCAATCGATGGGAACACAGAGGATG 430
QY 422 ACCCTACATGCTCAGATGTTCGCGAGGTCTTCGCTGACGAGAACCCCGCAGTACG 481
DB 431 ATGCTCATATGATGAGATGTTTCAAGACGTGTTCACGGTGAAGATGTAGGCACCTACG 490
QY 482 TGTGGGAGCGCAACTPACCAACAGCCCTGGGAACCCGCGCGCGCGCGCTGGCGCGC 541
DB 491 TATGGAAGGGAATTTACCACTGTTTAAAGCCTGAGAGAACCAATAGGGATCTCTGTC 550
QY 542 GAGCCTATCGGAGGTGAGGCGGAGGATCTTGGCGGCTGCGCATGAGGCGCTGTGTA 601
DB 551 CTGAGGTTTCTAAGCCCTTGAATGAAGATGAAGGTATCAATATGTTGGAAGAACTTGTTA 610
QY 602 GGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGCGAGTCTATGACA 661
DB 611 GAAACAAGAGATGTGTGTGTCATATGCAACAGTGGGCGGGAAGAGAGGTCTATGACA 670
QY 662 CAGCCTTCAAGCGCTCCGAGATTCCGTGGACTACGTGACGTCTGTGACTCGGACACA 721
DB 671 CAGCATTCAGGCCATTGGGACTTCTGTGACTATGTACAGTCTGTGACTTCGGACACA 730
QY 722 GGTGGACCCCATGCACTGCTGGAGCTCGTGGGTTACTGGGCTACTGGACGAGACCCCGGGTAG 781
DB 731 AACTGGATGAATGGCAACAGTGGAAATGGTGAAGGTTCTGGAAATCAATGACATGTACG 790
QY 782 GGGCTGTGTGGGACGTGGGATCTTAACCTCTGGACTCTCTGGGTGAGTCTCTTAA 841
DB 791 GCGCAGTGGGAGGAGACGTTCCGATTTCTGAACCTTATGATTTCTTATTTAGTTTCTGA 850
QY 842 GCAGCTCCGATACTGGGTAGCTTCAATGTGGAGCGGCTTGTTCAGAGTACTTCCACT 901
DB 851 GCAGCTCCGTTACTGGATGGGTTTAACTGTGGAGAGGCGCTGCCAGTCTTACTTCGACT 910
QY 902 GTGTATCTTCGATCAGCGGCTCTCTAGAAATCTTGCCTCCAGGCGCCCGGAGGACCGGATGA 961
DB 911 CGGTGTCTGTATAAGTGAACCTCTGGGAATGTATACCGGAACAAATTTCTCCAGGTGTTT 970
```

```
QY 962 TG 963
DB 971 TG 972

RESULT 15
US-09-902-939-3
; Sequence 3, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; APPLICANT: Weiliam Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-US0
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: mouse
US-09-902-939-3

Query Match      21.8%; Score 236.4; DB 10; Length 1665;
Best Local Similarity 58.5%; Pred. No. 1.1e-48;
Matches 511; Conservative 0; Mismatches 321; Indels 42; Gaps 4;

QY 61 CGGAGGCTGCTGACCATCGCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
DB 28 CGTGTGTGGGCACCAAGTCTGTTTGGCTGTGTAGTGTCTGGAGGCAATCCTTGGCGCTAT 87
QY 121 GCCCGCGGGGTGCGCTGGGCTCCGATCGCTACCGGCTCTCGGCTTCTGGGCTCTACGGG 180
DB 88 GTGACAGGCTACCAAGTTTATCCACAGAAAAGCACTACCTGTCTTTGGGCTCTACGGT 147
QY 181 GCCTTCTCTTTCAGCGCACTGTGGCGCAGAGCTCTTCGCTACCTTGGAGCACCGGCGG 240
DB 148 GCCATCTTGGGTCTACATCTGCTCATCCAGAGCTGTGTTGCTTCTTGGAGCACCGTGA 207
QY 241 GTGCGCGCGCGCGCGCG-----GGGCGCGCTGTGATGACGCCACCGCGCGCAGTGTGGCG 294
DB 208 ATGCGCAGGCGAGCGCGCGCCCTCAAGTCACTGCTCCACAGAGTCTCGGTTCAAGTGA 267
QY 295 CTGACCATCTCCGCTTACAGAGGACCCGCGTACCTGCGCCAGTGTGCTGGCGTCCGCC 354
DB 268 CTCTGCAATGCTGCTTACCAAGAGGACCCCGAATACCTGCGCAAGTGTCTTCTCGCTCAGCT 327
QY 355 CGCGCCCTGCTGTACCGCGCGCGCGCTGCGCGCTCTCATGTGTGGATGGCAACCGC 414
DB 328 CAGCGCATGCTTCTTCAAAAC-----CTCAAGGTGTGTATGTGTGATGGCAATCGC 381
QY 415 GCCGAGGACCTCTACATGCTGACATGTTCCGAGGTCTTTCGCTGACGAGGACCCCGCC 474
DB 382 CAGGAAGATACCTACATGTTGGAATCTTCCATGAGGTGTGCTGGTGGCACTGAGCAGCT 441
QY 475 ACGTACTGTGGAGCGGCAACTACCAAGCCCTTGGGAAACCCGCGCGCGCGCGCGGTG 534
DB 442 GGCTTCTTGTGTGGCGTAGCAATTTCCAT-----GAGCGCGGTGAAGGA 486
QY 535 GGCGCCGAGCTTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAGCGG 594
DB 487 GAGACAGAGCGCAGCTTGCAGGAAGGATGGAGCGT-----GTCCGAGCT 531
QY 595 CTGTGTAGGACTCTGAGGTGCTGTGTGGCGCAGCGCTGGGCGCGCAAGCGCGAGGTC 654
DB 532 GTGTGTGGGCCAGCACTTCTCATGATCATGACAGAGTGGGGGGGCAAGCGTGAAGTC 591
QY 655 ATGTACACAGCTTTCAAGCGCTCGGAGATTCGGTGAAGTACGTGCAAGTCTGTGACTCG 714
DB 592 ATGTACACTGCTTCAAGGCCCTTGGCAACTCACTGAGCTACATCCAGGTTGTGACTCT 651
QY 715 GACACAGGTTGGACCCCATGCGACTGCTGAGGCTCGTGGGCTACTTGGACGAGGACCCC 774
```


THIS PAGE BLANK (USTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:46:32 ; Search time 101.25 Seconds
(without alignments)
1375.148 Million cell updates/sec

Title: US-10-672-399-6

Perfect score: 1864

Sequence: 1 MRQDAPKPTPAARCSGLA.....ALVLPAVAEPADTLVQVVLP 360

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980a:*

2: Geneseqp1990a:*

3: Geneseqp2000a:*

4: Geneseqp2001a:*

5: Geneseqp2002a:*

6: Geneseqp2003a:*

7: Geneseqp2003bs:*

8: Geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1605	86.1	578	2 AAW26765	AAW26765 Human hya
2	1512	81.1	582	3 AAY78135	AAY78135 Recombina
3	1512	81.1	583	2 AAW30704	AAW30704 Mouse hya
4	1512	81.1	583	2 AAY32503	AAY32503 Hyaluro
5	1512	81.1	583	3 AAY68491	AAY68491 Mouse hya
6	1512	81.1	583	3 AAB09948	AAB09948 Murine HA
7	1512	81.1	583	6 AABP96028	ABP96028 Mouse hya
8	1512	81.1	583	7 AAE339152	AAE339152 Mouse hya
9	1425	76.4	543	2 AAW36503	AAW36503 Human hya
10	1419.5	76.2	584	3 AAY78132	AAY78132 Recombina
11	1419	76.1	582	3 AAY78138	AAY78138 Recombina
12	1419	76.1	583	3 AAY78131	AAY78131 Recombina
13	1238	66.4	563	3 AAY78129	AAY78129 Recombina
14	1000.5	53.7	577	3 AAY78128	AAY78128 Recombina
15	992	53.2	573	3 AAY78127	AAY78127 Recombina
16	876.5	47.0	552	3 AAY78133	AAY78133 Recombina
17	876.5	47.0	553	3 AAY78136	AAY78136 Recombina
18	811	43.5	552	7 ADC49213	ADC49213 Rabbit hy
19	804.5	43.2	588	7 ADD93928	ADD93928 Xenopus i
20	800	42.9	554	2 AAW50010	AAW50010 Murine hy
21	800	42.9	554	3 AAY68493	AAY68493 Mouse hya
22	800	42.9	554	6 AABP96030	ABP96030 Mouse hya
23	800	42.9	554	7 AAE339154	AAE339154 Mouse hya
24	800	42.9	554	8 ADJ76214	ADJ76214 Marker Ge
25	790.5	42.4	552	3 AAY78140	AAY78140 Recombina

26	790.5	42.4	553	6 ABR48209	ABR48209 Human bla
27	790.5	42.4	553	6 ABUS6490	ABUS6490 Lung canc
28	790.5	42.4	553	6 ABUS6686	ABUS6686 Lung canc
29	790.5	42.4	553	7 ADN39052	ADN39052 Cancer/an
30	790.5	42.4	553	8 ADJ75425	ADJ75425 Marker ge
31	790.5	42.4	553	8 ADN05116	ADN05116 Antipso
32	787.5	42.2	552	3 AAY78142	AAY78142 Recombina
33	787.5	42.2	552	3 AAY78139	AAY78139 Recombina
34	784	42.1	553	3 AAY78134	AAY78134 Recombina
35	783.5	42.0	552	2 AAW50009	AAW50009 Murine hy
36	783.5	42.0	552	3 AAY78141	AAY78141 Recombina
37	783.5	42.0	552	3 AAY68492	AAY68492 Mouse hya
38	783.5	42.0	552	5 ABB57265	ABB57265 Mouse isc
39	783.5	42.0	552	6 AABP96029	ABP96029 Mouse hya
40	783.5	42.0	552	7 ADA49686	ADA49686 Mouse hya
41	783.5	42.0	552	7 ADA49684	ADA49684 Human hya
42	783.5	42.0	552	7 AAE339153	AAE339153 Mouse hya
43	783.5	42.0	552	7 ADC59297	ADC59297 Human HAS
44	783.5	42.0	552	7 ADC59299	ADC59299 Mouse HAS
45	783.5	42.0	552	7 ADD48818	ADD48818 Human Pro

ALIGNMENTS

RESULT 1
AAW26765
ID AAW26765 standard; protein; 578 AA.
XX AC AAW26765;
XX DT 21-MAY-1998 (first entry)
XX DE Human hyaluronan synthase.
XX KW Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
XX KW wound healing; vulnery; tissue repair; scar; keloid; therapy.
XX OS Homo sapiens.
XX FH Key
FT Binding-site Location/Qualifiers
FT Modified-site /note= "putative hyaluronan binding site"
FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"
FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"
FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"
FT Modified-site /note= "consensus phosphorylation sequence for cAMP-dependent protein kinase"
WO9740174-A1.
PD 30-OCT-1997.
XX PF 17-APR-1997; 97WO-US006350.
XX PR 22-APR-1996; 96US-00635552.
XX PA (LEUK-) LEUKOSITE INC.
XX PI Brieskin MJ;
XX DR WPI; 1997-549359/50.
XX DR N-PSDB; AAT99541.
XX PT Human hyaluronan synthase - useful for recombinant production of
PT hyaluronic acid for wound healing, tissue repair and reducing

PT hypertrophic scar and keloid formation.

PS Claim 5; Page 38-40; 58pp; English.

XX This protein comprises human hyaluronan synthase (HAS), an enzyme
CC involved in the synthesis of hyaluronan (hyaluronic acid) and which has
CC the ability to confer cell adhesion by the lymphocyte receptor CD44. The
CC amino acid sequence was deduced from clone 30C (see AAT99541), which was
CC obtained from a human mesenteric lymph node expression library using an
CC expression cloning system developed to isolate cDNA clones that encode
CC proteins that confer adhesion of the murine T cell lymphoma TK1. The
CC isolated clone can be utilised in a claimed method for producing HAS in
CC host cells. Such host cells are used in a claimed method for the
CC production of hyaluronan. Hyaluronan is useful for wound healing and
CC tissue repair, and can reduce or prevent hypertrophic scars and keloid
CC formation. It is also used in eye surgery as a replacement for vitreous
CC fluid

SQ Sequence 578 AA;

Query Match 86.1%; Score 1605; DB 2; Length 578;
Best Local Similarity 96.0%; Pred. No. 1.8e-170;
Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYG 60
DB 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYG 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA 120
QY 121 LLYPRARLVLMVDGNRAEDLYMVDMPREVFADPATYWDGNYHQPWEPAAAGVGA 180
DB 121 LLYPRARLVLMVDGNRAEDLYMVDMPREVFADPATYWDGNYHQPWEPAAAGVGA 180
QY 181 GAYREVEADPGRLAVEALVTRRCVCAQRWGKGKREVMYTAFAKALGDSYDVYVQVCSDT 240
DB 181 GAYREVEADPGRLAVEALVTRRCVCAQRWGKGKREVMYTAFAKALGDSYDVYVQVCSDT 240
QY 241 RLDPMALLELVRLVLEDDPRVGAAGVGDVRLNPLDSWVSFLSRLRYWVAFNVERACQSYFH 300
DB 241 RLDPMALLELVRLVLEDDPRVGAAGVGDVRLNPLDSWVSFLSRLRYWVAFNVERACQSYFH 300
QY 301 CVSCISGPLESCPGPREHAMPSFL 325
DB 301 CVSCISGPL-----GLYRNLLQQFL 321

RESULT 2

AAV78135
ID AAV78135 standard; protein; 582 AA.

AC AAV78135;

XX 27-APR-2000 (first entry)

DT Recombinant chimeric hyaluronate synthase modified protein #9.

DE Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.

XX Mus sp.

OS Synthetic.

OS Chimeric.

XX JP2000004886-A.

XX 11-JAN-2000.

XX 24-JUN-1998; 98JP-00193788.

XX 24-JUN-1998; 98JP-00193788.

XX (SEKG) SEIKAGAKU KOGYO CO LTD.

XX WPI; 2000-140125/13.

XX A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.

XX Claim 10; Page; 30pp; Japanese.

XX The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 593). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAV78127 to AAV78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim

XX Sequence 582 AA;

Query Match 81.1%; Score 1512; DB 3; Length 582;
Best Local Similarity 90.6%; Pred. No. 5.1e-160;
Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3;

QY 3 QQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYGAF 62
DB 2 QQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYGAF 61
QY 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILAS 117
DB 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILAS 121
QY 118 ARALLYPRARLVLMVDGNRAEDLYMVDMPREVFADPATYWDGNYHQPWEPAAAG 176
DB 122 ARALLYPHTRLVLMVDGNRAEDLYMVDMPREVFADPATYWDGNYHQPWEPAAATG 181
QY 177 AVGAGAREVEADPGRLAVEALVTRRCVCAQRWGKGKREVMYTAFAKALGDSYDVYVQVC 236
DB 182 AVGAGAREVEADPGRLAVEALVTRRCVCAQRWGKGKREVMYTAFAKALGDSYDVYVQVC 241
QY 237 DSDTRLDPMALLELVRLVLEDDPRVGAAGVGDVRLNPLDSWVSFLSRLRYWVAFNVERACQ 296
DB 242 DSDTRLDPMALLELVRLVLEDDPRVGAAGVGDVRLNPLDSWVSFLSRLRYWVAFNVERACQ 301
QY 297 SYFHCVSCISGPLESCPGPREHAMPSFL 325
DB 302 SYFHCVSCISGPL-----GLYRNLLQQFL 326

RESULT 3

AAW30704

ID AAW30704 standard; protein; 583 AA.

XX AAW30704;

XX 17-OCT-2003 (revised)

DT 19-DEC-1997 (first entry)

XX Mouse hyaluronate synthase enzyme.

XX Hyaluronate synthase; hyaluronic acid; polysaccharide.

XX Mus sp; (strain FM3A).

XX JP09224674-A.

XX 02-SEP-1997.

```

XX Mus sp.
OS JPI1196875-A.
XX
XX PD 27-JUL-1999.
XX
XX PF 14-JAN-1998; 98JP-00006191.
XX
XX PR 14-JAN-1998; 98JP-00006191.
XX
XX PA (SEKG ) SEIKAGAKU KOGYO CO LTD.
XX
XX DR WPI; 1999-496653/42.
XX
XX DR N-PSDB; AAZ10862.
XX
XX PT New promoter DNA of hyaluronate synthase - used to specifically express
XX gene with cell growth inhibiting activity.
XX
XX PS Example 1; Page 8-10; 13pp; Japanese.
XX
XX CC This sequence is encoded by the mouse hyaluronate synthase gene, which
XX was used to isolate the hyaluronate synthase promoter of the invention.
XX CC The promoter can be used for specifically expressing a gene having cell
XX growth inhibiting activity
XX
XX SQ Sequence 583 AA;

      Query Match          81.1%; Score 1512; DB 2; Length 583;
Best Local Similarity    90.6%; Pred. No.5.le-160;
Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3

Qy   3 QQDAPKPTPAARRCCGLARRVLTITAFALLIIGLMTWYAAGVPLASDRYGLLAFLGYGAF 62
Db   2 RQDMPKPEEARCCSGLARRALTIIFALLIIGLMTWYAAGVPLASDRYGLLAFLGYGAF 61
Qy   63 LSAHLVAQSILPAYLEHRRVAAAAAR-----GPLDAATARSVALTISAYOEDPAYLRQCCLAS 117
Db   62 LSAHLVAQSLPAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYOEDPAYLRQCCLTS 121
Qy   118 ARALLYPRRLRVLMVDGNRAEADLYMDMREVFPADEDPATYWDGNVHPWEPA-AAAG 176
Db   122 ARALLYPHTLRVLVMVDGNRAEADLYMDMREVFPADEDPATYWDGNVHPWEPAEATG 181
Qy   177 AVGAGAVREVEAEPGRILAVALVRTRCVCVAORWGCKREVMTYAFKALGSDVDYVQC 236
Db   182 AVGEGAVREVEAEPGRILAVALVRTRCVCVAQRWGCKREVMTYAFKALGSDVDYVQC 241
Qy   237 DSDTRLDFMALLELVRLVLDPRVGAVGGDVRIINPLDSWVSFSSLRYYWAFNVERAQ 296
Db   242 DSDTRLDFMALLELVRLVLDPRVGAVGGDVRIINPLDSWVSFSSLRYYWAFNVERAQ 301
Qy   297 SYFHCVCSISGPSCPGPREHAMMPSEL 325
Db   302 SYFHCVCSISGPL----GLYRNLLQQPL 326

RESULT 5
AAy68491
ID   AAY68491 standard; protein; 583 AA.
XX
XX AC AAY68491;
XX
XX DT 27-APR-2000 (first entry)
XX
XX DE Mouse hyaluronate synthase modified protein HAS1.
XX
XX KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
XX research reagent; biochemical research; medical development.
XX
XX OS Mus sp.
XX
XX FN JP2000004886-A.
```


Db	242	DSDTRLDPMALLELVRLVDEDRVGVAGGDDVRILNPLNPSQSVSFLSSLRYYWAFNVERACQ	301
Qy	297	SYFHCVSCISGPLESCPGPREHAMMPSFL	325
Db	302	SYFHCVSCISGPL---GLYRNLLQQFL	326
Db	302	SYFHCVSCISGPL---GLYRNLLQQFL	326
RESULT 8			
AAE39152	ID	AAE39152 standard; protein; 583 AA.	
XX	AC	AAE39152;	
XX	AC	AAE39152;	
XX	DT	18-DEC-2003 (first entry)	
XX	DE	Mouse hyaluronan synthase (HAS) 1.	
XX	KW	Mouse; hyaluronic acid; HA; hyaluronan synthase; HAS; dry eye syndrome;	
XX	KW	eye cell; osteoarthritis; gene therapy; enzyme.	
XX	OS	Mus sp.	
XX	PN	US2003087850-A1.	
XX	PD	08-MAY-2003.	
XX	PF	10-JUL-2001; 2001US-00902939.	
XX	PR	10-JUL-2001; 2001US-00902939.	
XX	PA	(DEHA)/ DEHAZYA P.	
XX	PA	(CHEN)/ CHEN W.	
XX	PI	Dehazya P, Chen W;	
XX	DR	WPI; 2003-755151/71.	
XX	DR	N-PSDB; AAD59442.	
XX	PT	Dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate	
XX	PT	comprising derivatized hyaluronic acid cross linked to nucleic acid	
XX	PT	encoding hyaluronan synthase useful for treating dry eye syndrome.	
XX	PS	Claim 18; Page 15-16; 31pp; English.	
XX	CC	The invention relates to dihydrazide derivatised hyaluronic acid (HA) /	
XX	CC	nucleic acid bioconjugate comprising derivatised HA cross linked to	
XX	CC	nucleic acid encoding hyaluronan synthase (HAS). The invention is useful	
XX	CC	for transfected an eye cell of an individual. It is useful for treating	
XX	CC	dry eye syndrome and osteoarthritis of the particular joints. The	
XX	CC	invention is also useful in gene therapy. The present sequence is mouse	
XX	CC	HAS1 enzyme	
XX	SQ	Sequence 583 AA;	
Qy	3	QDAPKPTPAARCCSGLARVLTITAFALLIIGLMTWAYAAGVPLASDRYGLLAFGLYGAF	62
Db	2	QDMPKPSEARCCSGLARALTIIFALLIIGLMTWAYAAGVPLASDRYGLLAFGLYGAF	61
Qy	63	LSAHLVAQSLPAYLEHRRVAAAAR- - - -GPLDAATARSVALTISAYQEDPAYLRQCLAS	117
Db	62	LSAHLVAQSLPAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCLTS	121
Qy	118	ARALLYPRLRLVLMVVDGNRAEDLYVMDFREVFADDPATYYVMDGNHOFWEPA-AAG	176
Db	122	ARALLYPHTRLRLVLMVVDGNRAEDLYVMDFREVFADDPATYYVMDGNHOFWEPAEATG	181
Qy	177	AVGAGAYREVEADPGRLAVEALVTRRCVCVAORWGKREVMYTAFAKALGSDVDYQVC	236
Db	182	AVGAGAYREVEADPGRLAVEALVTRRCVCVAORWGKREVMYTAFAKALGSDVDYQVC	241

[illegible]

Db 122 ARALLYPHTRLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHYQWEPAEATG 181
QY 177 AVGAGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYT-APKALGSDVDYQV 235
Db 182 AVGAGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGNSVDYIQV 241
QY 236 CSDTRLDPMALLELVRLVLDDEPRVAVGVDVRLNPLDSWVSFLSSLRVYVAFNVERAC 295
Db 242 CSDTVLPDPACTIEMRLVLEEDPQVGGVGDVQILNKYDSWISFLSSVRYWMAFNVERAC 301
QY 296 QSYFHCVCISGPLESCPCPREHAMMPSFL 325
Db 302 QSYFGVCQCISGPL-----GMYRNSLLQQL 327

RESULT 11
AA78138
ID AAY78138 standard; protein; 582 AA.
XX
AC AAY78138;
DT 27-APR-2000 (first entry)
XX
DE Recombinant chimeric hyaluronate synthase modified protein #12.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX JP2000004886-A.
XX
PD 11-JAN-2000.
XX
PF 24-JUN-1998; 98JP-00193788.
XX
PR 24-JUN-1998; 98JP-00193788.
XX
PA (SEK) SEIKAGAKU KOGYO CO LTD.
XX
DR WPI; 2000-140125/13.
XX
PT A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
PS Claim 13; Page; 30pp; Japanese.

CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAY78127 to AAY78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim
XX
SQ Sequence 582 AA;

Query Match 76.1%; Score 1419; DB 3; Length 582;
Best Local Similarity 83.9%; Pred. No. 1.4e-149;
Matches 276; Conservative 18; Mismatches 25; Indels 10; Gaps 3;

QY 3 QODAPKPTFAARRCCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLYGAF 62
Db 2 RQDMKPSEARCCSGLARRALTITFALLILGLMTWAYAAGVPLASDRYGLLAFLYGAF 61

QY 63 LSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCLAS 117
Db 62 LSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCLTS 121
QY 118 ARALLYPARLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHYQWEPAEATG 176
Db 122 ARALLYPHTRLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHYQWEPAEATG 181
QY 177 AVGAGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFKALGSDVDYQV 236
Db 182 AVGAGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGSDVDYQV 241
QY 237 DSDTRLDPMALLELVRLVLDDEPRVAVGVDVRLNPLDSWVSFLSSLRVYVAFNVERAC 296
Db 242 DSDTMLDPASSVMVKVLEEDPMVGGVGDVQILNKYDSWISFLSSVRYWMAFNTERAC 301
QY 297 SYFHCVCISGPLESCPCPREHAMMPSFL 325
Db 302 SYFGVCQCISGPL-----GMYRNSLLHEFV 326

RESULT 12
AA78131
ID AAY78131 standard; protein; 583 AA.
XX
AC AAY78131;
DT 27-APR-2000 (first entry)
XX
DE Recombinant chimeric hyaluronate synthase modified protein #5.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX JP2000004886-A.
XX
PD 11-JAN-2000.
XX
PF 24-JUN-1998; 98JP-00193788.
XX
PR 24-JUN-1998; 98JP-00193788.
XX
PA (SEK) SEIKAGAKU KOGYO CO LTD.
XX
DR WPI; 2000-140125/13.
XX
PT A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
PS Claim 6; Page; 30pp; Japanese.

CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAY78127 to AAY78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim
XX
SQ Sequence 583 AA;

Query Match 76.1%; Score 1419; DB 3; Length 583;
Best Local Similarity 83.9%; Pred. No. 1.4e-149;
Matches 276; Conservative 18; Mismatches 25; Indels 10; Gaps 3;

THIS PAGE BLANK (138710)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:06:48 ; Search time 21.1364 Seconds
(without alignments)
1638.789 Million cell updates/sec

Title: US-10-672-399-6
Perfect score: 1864
Sequence: 1 MRQDAPKPTPAARCSGLA.....ALVLPAAVEADTLVQVVLP 360

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1429	76.7	543	2 JC4812	hyaluronan synthas
2	804.5	43.2	588	2 A43740	DG42 protein - Afr
3	310	16.6	568	2 T17588	hyaluronoglycosami
4	214	11.5	424	1 Z2ZRCU	nodulation protein
5	204.5	11.0	426	2 A95321	NodC N-ACETYLGLUCO
6	201.5	10.8	426	1 Z2ZRC4	nodulation protein
7	194.5	10.4	413	1 S34305	nodulation protein
8	173.5	9.3	424	1 S12793	nodulation protein
9	163.5	8.8	1086	2 JC6079	chitin synthase (E
10	159	8.5	1195	2 S61886	chitin synthase (E
11	158	8.5	395	1 JQ0396	nodulation protein
12	156.5	8.4	1165	1 S45879	chitin synthase (E
13	151.5	8.1	419	2 A53100	hyaluronate synth
14	151.5	8.1	1112	2 T30202	probable chitin sy
15	149.5	8.0	395	2 A48755	hyaluronan synthas
16	149.5	8.0	1239	2 T42020	class IV chitin sy
17	145.5	7.8	1041	2 T31037	chitin synthase (E
18	142	7.6	1175	2 S39951	chitin synthase (E
19	137.5	7.4	365	2 E59102	hypothetical prote
20	135	7.2	428	1 E38180	nodulation protein
21	123.5	6.6	1103	2 T42022	probable chitin sy
22	117	6.3	743	2 T34632	probable bi-functi
23	116	6.2	447	2 A97211	glycosyltransferas
24	114.5	6.1	1498	2 S78102	chitin synthase (E
25	114.5	6.1	1852	2 JC5546	chitin synthase (E
26	114.5	6.1	1869	2 A59290	class V chitin syn
27	111.5	6.0	412	2 B90075	intercellular adhe
28	108.5	5.8	869	2 H83500	probable glucosyl
29	105	5.6	615	2 E70663	probable PPE prote

30 104.5 5.6 3519 2 S43048 polyketide synthas
31 101.5 5.4 1198 2 T28678 polyketide synthas
32 100.5 5.4 3034 2 T14119 seven-pass transme
33 100 5.4 1155 2 AC2675 chromosome segrega
34 100 5.4 1165 2 A37457 structural mainten
35 99.5 5.3 291 2 B82645 phenylacetalddehyde
36 99.5 5.3 412 2 S77608 probable intercell
37 99 5.3 352 2 F75099 rhannosyl transfer
38 98.5 5.3 270 2 A82313 probable beta-keto
39 97 5.2 414 2 S18962 BBF15 protein - St
40 97 5.2 482 2 G83928 hypothetical prote
41 97 5.2 871 2 AE3085 two component sens
42 97 5.2 880 2 D98201 hypothetical prote
43 97 5.2 3573 2 S23070 erythronolide synt
44 96.5 5.2 420 2 D69769 cellulose synthase
45 96.5 5.2 460 2 H87396 conserved hypothet

ALIGNMENTS

RESULT 1

JC4812
hyaluronan synthase (EC 2.4.1.1-) - human
C:Species: Homo sapiens (man)
C>Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 02-Mar-2001
C:Accession: JC4812
R:Itano, N.; Kimata, K.
Biochem. Biophys. Res. Commun. 222, 816-820, 1996
A:Title: Molecular cloning of human hyaluronan synthase.
A:Reference number: JC4812; MUID:96244584; PMID:8651928
A:Accession: JC4812
A:Molecule type: mRNA
A:Residues: 1-543 <I>A>
A:Cross-references: DDBJ:D84424; NID:gl401033; PIDN:BAAL12351.1; PID:d1013030; PID:gl401030
C:Comment: This enzyme synthesizes a high molecular weight glycosaminoglycan composed of
C:Genetics:
A:Gene: GDB:HAS1; HAS
A:Cross-references: GDB:I220109; OMIM:601463
A:Map position: 19q13.4-19q13.4
C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein
F:58,167,308,321/Binding site: phosphate (Thr) (covalent) #status predicted
F:82,247/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 76.7%; Score 1429; DB 2; Length 543;
Best Local Similarity 95.2%; Pred. No. 4.8e-114; Mismatches 6; Indels 4; Gaps 1;
Matches 276; Conservative 4; Mismatches 6; Indels 4; Gaps 1;
Qy 36 MTWAYAGVPLASDRYGLLAFLGFLYGLAFSLAHLVAQSLFAYLEHRRVAAAARGPLDAATAR 95
Db 1 MTWAYAGVPLASDRYGLLAFLGFLYGLAFSLAHLVAQSLFAYLEHRRVAAAARGPLDAATAR 60
Qy 96 SVALTISAYQEDPAYLRQCCLASARALLYPRARLVLMVVDGNRAEDLVMDMFREVFADE 155
Db 61 SVALTISAYQEDPAYLRQCCLASARALLYPRARLVLMVVDGNRAEDLVMDMFREVFADE 120
Qy 156 DPATVWDGNYHQWEPAAAGVAGAGAYREAEADPGRLAVEALVRRVCVCVAQRWGK 215
Db 121 DPATVWDGNYHQWEPAAAGVAGAGAYREAEADPGRLAVEALVRRVCVCVAQRWGK 180
Qy 216 REVMTYAFKALGDSVDYVQVCDSTRDLPALLLVRVLDPRVGVAGVGGDVRILNPLDS 275
Db 181 REVMTYAFKALGDSVDYVQVCDSTRDLPALLLVRVLDPRVGVAGVGGDVRILNPLDS 240
Qy 276 WVSFSLSLRYVWAFNVERACQSYFHCVSCISGPLESCPGPREHAMPSFL 325
Db 241 WVSFSLSLRYVWAFNVERACQSYFHCVSCISGPLESCPGPREHAMPSFL 286
RESULT 2
A43740
DG42 protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)

A:Reference number: JQ0393; MUID:90136519; PMID:2615763
A:Accession: JQ0396
A:Molecule type: DNA
A:Residues: 1-1395 <GOE>
A:Cross-references: UNIPROT:Q07755; GB:118897; NID:gl293899; PIDN:AAB51164.1; PID:g31029
A:Experimental source: strain ORS571
C:Comment: this is one of the proteins, coded by nodulation genes, that are required for
C:Genetics:
A:Gene: nodC
C:Superfamily: nodulation protein nodC
C:Keywords: nodulation

Query Match 8.5%; Score 158; DB 1; Length 395;
Best Local Similarity 26.1%; Pred. No. 9.4e-06;
Matches 87; Conservative 36; Mismatches 112; Indels 98; Gaps 16;

QY 49 DRYGLAFLGLYGAFLSAHLVAQSLFAYLEHRRVAAARGLDA--ATARSVALTISAYOE 106
DB 5 DVIGLLATAAAYVTLASAKVQV----FINVSVTDVAGLESALPLTPR-VDVIVPTFNE 59
QY 107 DPAYLRQCLASARALLYPRARLRVLMVVDGNRAE-----DLYMVDMPREVFADEDPAT 159
DB 60 NSTLLECVASICAQDY-RGPTIVVDDGSTNKTSTFHAVCDKYASDE-RFIFVELD--- 114
QY 160 VYWDGNHQPWPBPAAGAVGAGAYREVEADPGRLA-VEALVTRRCVVCVAQRWGGKREV 218
DB 115 -----QNKGTAAQMEAIRT----- 129
QY 219 MYTAFKALGDSVDYVQVCDSTRDPMALLELRVLDDEPRVAGGVGVRLNPLDSWVS 278
DB 130 -----DGDLLNVDSDTVIDKVVTKLASSM-RAPNVGVMGQLVAKNRSMILT 178
QY 279 FLSSLRVYVAFNVERACOSYFHCVSCISGPLESFCPPREHAMVP-----SELA-PVQ 329
DB 179 RLIDNEYWLACNEERIAQSRFGSVWCCCGP---CAMYRSALTPLLAEYEHOTFLGRPSN 235
QY 330 V-----HLQVPLL---LRDALVLPVAPVADTIV 355
DB 236 FGEDRHLTILMLKAGFRTGYVPSAVAR---TLV 265

RESULT 12
S45879
chitin synthase (EC 2.4.1.16) 3 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YBR023c; protein YBR0305
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S45879; S46554; S22776; A39639; S17247
R:Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45875
A:Accession: S45879
A:Molecule type: DNA
A:Residues: 1-1165 <GRI>
A:Cross-references: UNIPROT:P29465; EMBL:235892; NID:g536229; PIDN:CAA84965.1; PID:g5362
A:Experimental source: strain S288C
R:Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.
Yeast 10(Suppl.A), S75-S80, 1994
A:Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II fro
ly identified genes and a homologue of the SCO1 gene.
A:Reference number: S46551; MUID:94378725; PMID:8091864
A:Accession: S46554
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1165 <SMI>
A:Cross-references: EMBL:X76078; NID:g498748; PIDN:CAA53680.1; PID:g498752
A:Experimental source: strain S288C
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
R:Bulawa, C.E.
Mol. Cell. Biol. 12, 1764-1776, 1992
A:Title: CS2D, CS3, and CS4, genes required for chitin synthesis in Saccharomyces cere
blum species and Xenopus laevis.
A:Reference number: S22776; MUID:92195323; PMID:1532231

A:Accession: S22776
A:Molecule type: DNA
A:Residues: 1-1162, 'L', 1164-1165 <BUL>
A:Cross-references: EMBL:M73697; NID:gl72103; PIDN:AAA34844.1; PID:gl72104
R:Valdivieso, M.H.; Mol, P.C.; Shaw, J.A.; Cabib, E.; Duran, A.
J. Cell Biol. 114, 101-109, 1991
A:Title: CAL1, a gene required for activity of chitin synthase 3 in Saccharomyces cerevi
A:Reference number: A39639; MUID:91268144; PMID:2050737
A:Accession: A39639
A:Molecule type: DNA
A:Residues: 67-1165 <VAL>
A:Cross-references: GB:X57300; NID:g3359; PIDN:CAA40559.1; PID:g3360
C:Genetics:
A:Gene: SGD:CHS3; CAL1; CS2D; MIPS:YBR023c
A:Cross-references: MIPS:YBR023c; SGD:S0000227
A:Map position: 2R
C:Function:
A>Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosa
C:Superfamily: chitin synthase chs4
C:Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase; transmembrane
F:203-219/Domain: transmembrane #status predicted <TM1>
F:457-473/Domain: transmembrane #status predicted <TM2>
F:1018-1034/Domain: transmembrane #status predicted <TM3>
F:1035-1054/Domain: transmembrane #status predicted <TM4>
F:1060-1076/Domain: transmembrane #status predicted <TM5>
F:1084-1100/Domain: transmembrane #status predicted <TM6>

Query Match 8.4%; Score 156.5; DB 1; Length 1165;
Best Local Similarity 22.2%; Pred. No. 4.5e-05;
Matches 58; Conservative 48; Mismatches 88; Indels 67; Gaps 9;

QY 101 ISAVQEDPAYLRQCLASARALLYPRARLRVLMVVDG---NRAED-----LYMVDMPRE 150
DB 658 VTCYSEDEGLRTLDSLTSTDYPSNHLKLVVDCGLKSGNDKTTPEIALGWMDDFTV 717
QY 151 VFAEDPATYVWDGNHQPWPBPAAGA-----VGAGAYR-----EVAEDPGRLAVAL 199
DB 718 PPDEVKPYSYV-----AVASGSKRHNMAKIYAGFYKYDDSTIPPENQORVPIITI 767
QY 200 VTRRCVCAQRWG-----GKR--EVMTATF-----KAL 226
DB 768 V---KCGTPAEQGAAPGNRGRDSQIILMSFLEKITFDERTWQLEFOLLKNWIQTGLM 824
QY 227 GDSVDYVQVCDSTRDPMALLELRVLDDEPRVAGGVGVRLNPLDSWVSFLSLRYW 286
DB 825 ADFVETVLMVDADIKVFPDALTHVAVEMVKDPLINGLGCETKIANKAQSWTAIQVFYY 884
QY 287 VAFNVERACOSYFHCVSCISG 307
DB 885 ISHQAKAFESVFGSVTCCLPG 905

RESULT 13
A53100
hyaluronate synthase A (Hasa) - Streptococcus sp. (group A)
C:Species: Streptococcus sp.
C>Date: 06-Oct-1994 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: A53100
A:Title: Molecular characterization of hasa from an operon required for hyaluronic acid
J. Biol. Chem. 269, 169-175, 1994
A:Reference number: A53100; MUID:94103204; PMID:8276791
A:Accession: A53100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <DOU>
A:Cross-references: UNIPROT:Q8NKK1
A>Note: sequence extracted from NCBI backbone (NCBIN:141683, NCBI:141684)
C:Superfamily: nodulation protein nodC

Query Match 8.1%; Score 151.5; DB 2; Length 419;
Best Local Similarity 24.0%; Pred. No. 3.6e-05;
Matches 80; Conservative 38; Mismatches 116; Indels 99; Gaps 13;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:48:17 ; Search time 94.0909 Seconds
(without alignments)
1599.259 Million cell updates/sec

Title: US-10-672-399-6

Perfect score: 1864

Sequence: 1 MRQDAPKPTPAARRCSGLA.....ALVLPVAEPAADTLVQVVLP 360

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	1605	86.1	578	1 HAS1 HUMAN
2	1597	85.7	578	Q9ns49 homo sapien
3	1588	85.2	577	Q8lyh3 homo sapien
4	1545.5	82.9	583	Q6s742 papio anubi
5	1512	81.1	583	1 HAS1 MOUSE
6	1507	80.8	583	Q8CH93
7	929.5	49.9	458	Q6T488
8	811	43.5	552	Q95M29
9	808	43.3	557	Q6W9J2
10	804.5	43.2	588	1 HAS1 XENLA
11	803.5	43.1	554	Q9DG30
12	800.5	42.9	553	Q75R37
13	800	42.9	554	1 HAS3 MOUSE
14	798	42.8	554	Q8CEB9
15	798	42.8	554	Q8CH92
16	790.5	42.4	553	1 HAS3 HUMAN
17	790.5	42.4	553	Q96RV2
18	784.5	42.1	552	Q95M16
19	783.5	42.0	552	1 HAS2 HUMAN
20	783.5	42.0	552	1 HAS2 MOUSE
21	783.5	42.0	552	Q8H2J3
22	782.5	42.0	552	Q8S070
23	781.5	41.9	552	1 HAS2 BOVIN
24	780	41.8	552	Q9DG41
25	778.5	41.8	552	1 HAS2 RAT
26	771	41.4	552	1 HAS2 CHICK
27	750.5	40.3	551	1 HAS2 XENLA
28	739.5	39.7	583	1 HAS3 XENLA
29	738.5	39.6	583	Q8AZ10
30	592.5	31.8	281	2 Q8WT20
31	560.5	30.1	131	2 Q8BPN0

32 541 29.0 393 2 Q811Y6
33 351 18.8 156 2 Q18792
34 310 16.6 568 2 Q84419
35 310 16.6 568 2 Q9WFS9
36 310 16.6 568 2 Q9WFT0
37 291 15.6 134 2 Q90489
38 272 14.6 245 2 Q9GK14
39 214.5 11.5 452 2 Q9AQ23
40 214 11.5 424 1 NODC_RHILV
41 204.5 11.0 426 1 NODC_RHIME
42 201 10.8 452 2 Q8EX51
43 198.5 10.6 408 2 Q8PTX8
44 196.5 10.5 413 1 NODC_RHISN
45 195 10.5 443 2 Q8KLG3

ALIGNMENTS

RESULT 1

HAS1_HUMAN STANDARD; PRT; 578 AA.
AC Q92839; Q14470;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DN Hyaluronidic acid synthase 1 (HA synthase 1) (HuHAS1).
GN Name=HAS1; Synonyms=HAS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RX MEDLINE=96394438; PubMed=8798544; DOI=10.1074/jbc.271.38.23395;
RA Shyjan A.M., Heidlin P., Butcher E.C., Yoshino T., Brieskin M.J.;
RT "Functional cloning of the cDNA for a human hyaluronan synthase.";
RL J. Biol. Chem. 271:23395-23399(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96244584; PubMed=8651928; DOI=10.1006/bbrc.1996.0827;
RA Itano N., Kinata K.;
RT "Molecular cloning of human hyaluronan synthase.";
RL Biochem. Biophys. Res. Commun. 222:816-820(1996).
CC -|- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA) synthesis. Also able to catalyze the synthesis of chito-oligosaccharide depending on the substrate.
CC -|- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-glucuronosyl(1->3)](n) + 2n UDP.
CC -|- COFACTOR: Magnesium.
CC -|- PATHWAY: Hyaluronate synthesis.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -|- TISSUE SPECIFICITY: Highly expressed in ovary followed by spleen, thymus, prostate, testes and large intestine. Weakly expressed in small intestine.
CC -|- SIMILARITY: Belongs to the nodC/HAS family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; U59269; AAC50706.1; --
EMBL; D84424; BAA12351.1; ALT_INIT.
Genew; HGNC:4819; HAS1.
MIM; 601463; --

```

DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR GO; GO:0006024; P:glycosaminoglycan biosynthesis; TAS.
DR InterPro; IPR001173; Glyco trans 2.
DR Pfam; PF00535; Glycos transf 2; 1.
KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
FT DOMAIN 1 25 Cytoplasmic (Potential).
FT TRANSMEM 26 46 1 (Potential).
FT DOMAIN 47 52 Extracellular (Potential).
FT TRANSMEM 53 73 2 (Potential).
FT DOMAIN 74 399 Cytoplasmic (Potential).
FT TRANSMEM 400 420 3 (Potential).
FT DOMAIN 421 430 Extracellular (Potential).
FT TRANSMEM 431 451 4 (Potential).
FT DOMAIN 452 457 Cytoplasmic (Potential).
FT TRANSMEM 458 478 5 (Potential).
FT DOMAIN 479 497 Extracellular (Potential).
FT TRANSMEM 498 518 6 (Potential).
FT DOMAIN 519 540 Cytoplasmic (Potential).
FT TRANSMEM 541 561 7 (Potential).
FT DOMAIN 562 578 Extracellular (Potential).
FT CONFLICT 1 2 MR -> RS (in Ref. 2).
FT CONFLICT 34 34 G -> A (in Ref. 2).
SQ SEQUENCE 578 AA; 64884 MW; 355FD54B0899E43C CRC64;

Query Match 86.1%; Score 1605; DB 1; Length 578;
Best Local Similarity 96.0%; Pred. No. 4e-126;
Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 1 MRQODAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL 60
DB 1 MRQODAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL 60

QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120

QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMEFVFADEDPATYWDGNYHQWPEPAAAGAVGA 180
DB 121 LLYPRARLRVLMVDGNRAEDLYMVDMEFVFADEDPATYWDGNYHQWPEPAAAGAVGA 180

QY 181 GAYREVEAEDPGRLAVEALVTRRCVCAQWGGKREVMYTFAPKALGDSVDYVQVCDSDT 240
DB 181 GAYREVEAEDPGRLAVEALVTRRCVCAQWGGKREVMYTFAPKALGDSVDYVQVCDSDT 240

QY 241 RLDPMALLELRVLDDEPRVGAAGDVRLNPLDSWVSFLSSRLRYWYAFNVERACQSYFH 300
DB 241 RLDPMALLELRVLDDEPRVGAAGDVRLNPLDSWVSFLSSRLRYWYAFNVERACQSYFH 300

QY 301 CVSCISGPLESCPGPREHAMMPSFL 325
DB 301 CVSCISGPL-----GLYRNLLQQFL 321

RESULT 2
Q9NS49 PRELIMINARY; PRT; 578 AA.
AC Q9NS49;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hyaluronan synthase 1.
GN Name=HAS1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E.;
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gaines J.,
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
Andreise T., Trankheim M., Attix C., Amico-Keller G., Coefield J.,
Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
Olisen A.S., Carrano A.V.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP DOE Joint Genome Institute;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP Lamerdin J.E.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC018755; AA87845.1; -.
DR InterPro; IPR001173; Glyco trans 2.
DR InterPro; IPR00276; GPCR_Rhodopsn.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; UNKNOWN 1.
DR PROSITE; PS00185; IPNS 1; UNKNOWN 1.
SQ SEQUENCE 578 AA; 64831 MW; 2FE3A44B0D5380FF CRC64;

Query Match 85.7%; Score 1597; DB 2; Length 578;
Best Local Similarity 95.7%; Pred. No. 1.9e-125;
Matches 311; Conservative 3; Mismatches 7; Indels 4; Gaps 1;

QY 1 MRQODAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL 60
DB 1 MRQODAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL 60

QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120

QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMEFVFADEDPATYWDGNYHQWPEPAAAGAVGA 180
DB 121 LLYPRARLRVLMVDGNRAEDLYMVDMEFVFADEDPATYWDGNYHQWPEPAAAGAVGA 180

QY 181 GAYREVEAEDPGRLAVEALVTRRCVCAQWGGKREVMYTFAPKALGDSVDYVQVCDSDT 240
DB 181 GAYREVEAEDPGRLAVEALVTRRCVCAQWGGKREVMYTFAPKALGDSVDYVQVCDSDT 240

QY 241 RLDPMALLELRVLDDEPRVGAAGDVRLNPLDSWVSFLSSRLRYWYAFNVERACQSYFH 300
DB 241 RLDPMALLELRVLDDEPRVGAAGDVRLNPLDSWVSFLSSRLRYWYAFNVERACQSYFH 300

QY 301 CVSCISGPLESCPGPREHAMMPSFL 325
DB 301 CVSCISGPL-----GLYRNLLQQFL 321

RESULT 3
Q81YH3 PRELIMINARY; PRT; 577 AA.
AC Q81YH3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hyaluronan synthase 1.
GN Name=HAS1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

```

RL	Submitted (NOV-2003) to the EMBL/GenBank/DBSJ databases.
DR	EMBL; AY463695; AAR25554.1; -
DR	InterPro; IPR001173; Glyco trans 2.
DR	InterPro; IPR000276; GPCR_Rhodopa
DR	InterPro; IPR002057; Isopen_N_synth.
DR	PROSITE; PS00237; G_PROTEIN_RECIP_FL_1; UNKNOWN_1.
DR	PROSITE; PS00185; IPNS_1; UNKNOWN_1.
SQ	SEQUENCE 583 AA; 65596 MW; EA47354C89346F94 CRC64;
Query Match 82.9%; Score 1545.5; DB 2; Length 583;	
Best Local Similarity 91.5%; Pred. No. 4e-121;	
Matches 302; Conservative 6; Mismatches 13; Indels 9; Gaps 2	
Qy	1 MRQODAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLYG 60
Db	1 MTQRDTFKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLYG 60
Qy	61 AFLSAHLVAQSLFAYLEHRRVA-----AAAGRPDLAATARSVALTISAYQEDPAVLRQCL 115
Db	61 AFLSAHLVAQSLFAYLEHRRVAAAARRAARGRLDAATARSVALTISAYQEDPAVLRQCL 120
Qy	116 ASARALLYPRARLVLMVDGNRAEDLVMDMFRVFPADEDPATVVDGNHYHQWPEAAA 175
Db	121 VSARALLYPRARLVLMVDGNRRFEDLVMDMFRVFPADEDPATVVDGNHYHQWPEAAV 180
Qy	176 GAVGAGAYREVEAEDPGRLEAVEALVTRRCVCVAQRWGKREVMYTAFAKALGDSVDYVQV 235
Db	181 GAVGVGAYREVEAEDPGRLEAVEALVTRRCVCVAQRWGKREVMYTAFAKALGDSVDYVQV 240
Qy	236 CDSSTRLDPMALLBELRVLDBDPVRGAVGGVDRILNPLDSDWVSFLSSLRVYVAFNVERAC 295
Db	241 CDSSTRLDPMALLBELVQLDBDPVRGAVGGVDRILNPLDSDWVSFLSSLRVYVAFNVERAC 300
Qy	296 QSYFHCVCISGPLESCGPFREHAMPSFL 325
Db	301 QSYFHCVCISGPLE-----GLYRNLLQQFL 326
RESULT 5	
ID	HAS1_MOUSE
AC	Q61647;
CD	30-MAY-2000 (Rel. 39, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DE	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE	(Hyaluronic acid synthase 1) (HA synthase 1).
GN	Name=Has1; Synonyms=Has;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=96215261; PubMed=8626618; DOI=10.1074/jbc.271.17.9875;
RA	Itano N.; Kimata K.;
RT	"Expression cloning and molecular characterization of HAS protein, a
RT	eukaryotic hyaluronan synthase";
RT	J. Biol. Chem. 271:9875-9878(1996).
RN	[2]
RP	MUTAGENESIS.
RX	MEDLINE=20085071; PubMed=10617644; DOI=10.1074/jbc.275.1.497;
RA	Yoshida M.; Itano N.; Yamada Y.; Kimata K.;
RT	"In vitro synthesis of hyaluronan by a single protein derived from
RT	mouse HAS1 gene and characterization of amino acid residues essential
RT	for the activity";
RT	J. Biol. Chem. 275:497-506(2000).
CC	-I- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
CC	synthesis. Also able to catalyze the synthesis of chito-
CC	oligosaccharide depending on the substrate.
CC	-I- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
CC	glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-
CC	glucuronosyl(1->3)](n) + 2n UDP.

CC -I- COFACTOR: Magnesium.
 CC -I- PATHWAY: Hyaluronate synthesis.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -I- SIMILARITY: Belongs to the nodC/HAS family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; D82964; BAA11654.1; -.
 DR MGI; 106590; Has1.
 DR InterPro; IPR001173; Glyco trans 2.
 DR Pfam; PF00535; Glycos transf 2; 1.
 DR KEGG; Glycosyltransferase; Multigene family; Transferase; Transmembrane.
 FT DOMAIN 1 24
 FT TRANSMEM 25 45
 FT DOMAIN 46 51
 FT TRANSMEM 52 72
 FT DOMAIN 73 404
 FT TRANSMEM 405 425
 FT DOMAIN 426 435
 FT TRANSMEM 436 456
 FT DOMAIN 457 462
 FT TRANSMEM 463 483
 FT DOMAIN 484 501
 FT TRANSMEM 502 522
 FT DOMAIN 523 545
 FT TRANSMEM 546 566
 FT DOMAIN 567 583
 FT TRANSMEM 584 81
 FT DOMAIN 81 84
 FT POLY-LEU. 516 519
 FT CARBOHYD 489 489
 FT MUTAGEN 242 242
 FT MUTAGEN 311 311
 FT MUTAGEN 312 312
 FT MUTAGEN 313 313
 FT MUTAGEN 314 314
 FT MUTAGEN 314 314
 FT MUTAGEN 344 344
 FT MUTAGEN 380 380
 FT MUTAGEN 383 383
 FT MUTAGEN 384 384
 FT SEQUENCE 583 AA; 65544 MW; 7AF9273E7B314728 CRC64;
 Query Match 81.1%; Score 1512; DB 1; Length 583;
 Best Local Similarity 90.6%; Pred. No. 2.6e-118;
 Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3;
 QY 3 QDAPKPTPAARCSGLARRVLTIFALLILGLMTWAYAAGVPLASDRYGLLAFLGLYCAP 62
 DB 2 RQDMKPESEARCCSGLARRVLTIFALLILGLMTWAYAAGVPLASDRYGLLAFLGLYCAP 61
 QY 63 LSAHLVAQSLFAYLEHRRVAAAAR-----GFLDAATARSVALTISAYQEDPAYLRQCCLAS 117
 DB 62 LSAHLVAQSLFAYLEHRRVAAAARSLAKGFLDAATARSVALTISAYQEDPAYLRQCCLTS 121
 QY 118 ARALLYPRALRLVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQPWEPAAAG 176
 DB 122 ARALLYPHTRLRLVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQPWEPAAATG 181
 QY 177 AVGAGAYREVAEDPGRLAVALVTRRCVCVAQRWGCKREVMTAFKALGDSVDYVQVC 236
 DB 182 AVGAGAYREVAEDPGRLAVALVTRRCVCVAQRWGCKREVMTAFKALGDSVDYVQVC 241
 QY 237 DSDTRLDPMALLELRVLDEDPKRGVGGVDRILNPLDSWVSFLSSLRVYVAFNVERACQ 296
 DB 242 DSDTRLDPMALLELRVLDEDPKRGVGGVDRILNPLDSWVSFLSSLRVYVAFNVERACQ 301

QY 297 SYFHCVSCISGPLESCPGPREHAMPSFL 325
 DB 302 SYFHCVSCISGPLESCPGPREHAMPSFL 326
 RESULT 6
 Q8CH93 PRELIMINARY; PRT; 583 AA.
 AC Q8CH93;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hyaluronan synthase 1.
 GN Name=HAS1;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14724275; DOI=10.1074/jbc.M313178200;
 RA Itano N., Sawai T., Atsumi F., Miyaishi O., Taniguchi S., Kannagi R.,
 RA Hamaguchi M., Kimata K.;
 RT "Selective expression and functional characteristics of three
 RT Mammalian hyaluronan synthases in oncogenic malignant
 RT transformation.";
 RL J. Biol. Chem. 279:18679-18687 (2004).
 DR EMBL; AB097568; BAC43730.1; -.
 DR InterPro; IPR001173; Glyco trans 2.
 DR InterPro; IPR02057; Isopen N synth.
 DR PROSITE; PS00185; IPNS 1; UNKNOWN 1.
 DR PROSITE; PS00185; IPNS 1; UNKNOWN 1.
 SQ SEQUENCE 583 AA; 65724 MW; 35513C6B21DE4B8D CRC64;
 Query Match 80.8%; Score 1507; DB 2; Length 583;
 Best Local Similarity 90.3%; Pred. No. 6.8e-118;
 Matches 297; Conservative 5; Mismatches 17; Indels 10; Gaps 3;
 QY 3 QDAPKPTPAARCSGLARRVLTIFALLILGLMTWAYAAGVPLASDRYGLLAFLGLYCAP 62
 DB 2 RQDMKPESEARCCSGLARRVLTIFALLILGLMTWAYAAGVPLASDRYGLLAFLGLYCAP 61
 QY 63 LSAHLVAQSLFAYLEHRRVAAAAR-----GFLDAATARSVALTISAYQEDPAYLRQCCLAS 117
 DB 62 LSAHLVAQSLFAYLEHRRVAAAARSLAKGFLDAATARSVALTISAYQEDPAYLRQCCLTS 121
 QY 118 ARALLYPRALRLVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQPWEPAAAG 176
 DB 122 ARALLYPHTRLRLVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQPWEPAAAG 181
 QY 177 AVGAGAYREVAEDPGRLAVALVTRRCVCVAQRWGCKREVMTAFKALGDSVDYVQVC 236
 DB 182 AVGAGAYREVAEDPGRLAVALVTRRCVCVAQRWGCKREVMTAFKALGDSVDYVQVC 241
 QY 237 DSDTRLDPMALLELRVLDEDPKRGVGGVDRILNPLDSWVSFLSSLRVYVAFNVERACQ 296
 DB 242 DSDTRLDPMALLELRVLDEDPKRGVGGVDRILNPLDSWVSFLSSLRVYVAFNVERACQ 301
 QY 297 SYFHCVSCISGPLESCPGPREHAMPSFL 325
 DB 302 SYFHCVSCISGPLESCPGPREHAMPSFL 326
 RESULT 7
 Q6T488 PRELIMINARY; PRT; 458 AA.
 ID Q6T488
 AC Q6T488;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hyaluronan synthase 1 (Fragment).
 GN Name=has1;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14729574;
 RA Bakkers J., Kramer C., Pothof J., Quaedvlieg N.B., Spaik H.P.,
 RA Hamerschmidt M.;
 RT "Has2 is required upstream of Rac1 to govern dorsal migration of
 RT lateral cells during zebrafish gastrulation.";
 RL Development 131:525-537(2004).
 DR EMBL; AY437407; AAR97372.1; --
 DR ZFIN; ZDB-GENE-040218-2; hasl.
 DR InterPro; IPR001173; Glyco_trans_2.
 FT NON TER 458 458
 SQ SEQUENCE 458 AA; 53188 MW; 4B421299ABBP9E6E CRC64;
 Query Match 49.9%; Score 929.5; DB 2; Length 458;
 Best Local Similarity 54.7%; Pred. No. 1.6e-69;
 Matches 174; Conservative 53; Mismatches 80; Indels 11; Gaps 4;
 QY 11 PAARCSGLARRVLTAFAALLILGLMTWAYAAGVPLASDRYGLLAFLGLYGAFLSAHLVAQ 70
 DB 5 PVLKFFSVLRAFLTLFALLVLGLWMLWAYIDGFOIATSPFGIISFGYGVLLSLJHLIQ 64
 QY 71 SLFAYLEHRRVAAARGPLDARTARSVALTISAYQEDPAYLRQCLASARALLYPRARLRV 130
 DB 65 SFFAPVEHGRMNARK--PCSYTKTIGTTISAYQEDPAYLRQCLASARALLYPRARLRV 121
 QY 131 LMVUDGNRAEDLYWDMFREFVADDPATYVWGNHYQFWPEPA--AGAVGAGAYREVE 187
 DB 122 VMVUDGNSEDDRYMLEMREFVADDPATYVWGNHYQFWPEPA--AGAVGAGAYREVE 180
 QY 188 AEDRCGLAVEALVTRRCVCAQRWGGKREVMYTAFAKLGSDVDYVQVCDSDTRLDPMAL 247
 DB 181 FEDPORLEVEIRTKKVCIMQKMGKREVMYTAFAKLGSDVDYVQVCDSDTRLDPMAL 240
 QY 248 LELVRLVDEPRVAGVGDVRIINPLDSWVSLSSRLRYVAFNVERACQSYFHCVSCISG 307
 DB 241 VELCKVLESQYKYGAVGVDVMIINLKDYSIFSMSSRLRYVAFNVERACQSYFHCVSCISG 300
 QY 308 PLESCPGPREHMMPSFL 325
 DB 301 PL----GLYRNDLLQOFL 314
 RESULT 8
 Q95MZ9 PRELIMINARY; PRT; 552 AA.
 AC Q95MZ9;
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Hyaluronic acid synthase 3.
 GN Name=HAS3;
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ohno S.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB055979; BAB63265.1; --
 DR InterPro; IPR001173; Glyco_trans_2.
 SQ SEQUENCE 552 AA; 63137 MW; 3CA805AD5347156 CRC64;
 Query Match 43.5%; Score 811; DB 2; Length 552;
 Best Local Similarity 54.6%; Pred. No. 1.8e-59;
 Matches 167; Conservative 38; Mismatches 83; Indels 18; Gaps 4;
 QY 21 RRLVTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLYGAFLSAHLVAQSLFAYLEHRR 80

DB 10 RVVGTSLPALAVLGGILAAAYVTGYQFIHTEKHYSFLGYLGLHLLIQSLFAFLEHRR 69
 QY 81 VAAARGPLDARTARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAE 140
 DB 70 MRRARRPLKPSRRRSVALCIAAYQEDPDYLRKLSRQRTAF--DLKVMVVDGNROE 127
 QY 141 DLYVDMFREFV--ADEDPATYVWGNHYQFWPEPAAGAVGAGAYREVEAEDPRLAVEAL 199
 DB 128 DAYMLDIFHEVLGGTEQAGFFWRSNPFHEAGEGETEASIQEGMER-----VRVAV 176
 QY 200 VTRRCVCAQRWGGKREVMYTAFAKLGSDVDYVQVCDSDTRLDPMALLELVRLVDEDP 259
 DB 177 VRTSTFCIMQKMGKREVMYTAFAKLGSDVDYVQVCDSDTVLDPACTIEMLRVLEEDPQ 236
 QY 260 VGAVGVDVRIINPLDSWVSLSSRLRYVAFNVERACQSYFHCVSCISGPLESCPGPREHA 319
 DB 237 VGGVGDVQIILNKYDWSIFLSSVRYWMAFNVERACQSYFGCVQCISGFL-----GMYRNS 292
 QY 320 MMSPSFL 325
 DB 293 LLQOFL 298
 RESULT 9
 Q6W9J2 PRELIMINARY; PRT; 557 AA.
 AC Q6W9J2;
 DT 05-JUL-2004 (TREMELrel. 27, Created)
 DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
 DE Hyaluronic acid synthase 3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14667843; DOI=10.1016/S0945-053X(03)00082-9;
 RA Vigetti D., Viola M., Gornati R., Ori M., Nardi I., Passi A.,
 RA De Luca G., Bernardini G.;
 RT "Molecular cloning, genomic organization and developmental expression
 RT of the xenopus laevis hyaluronan synthase 3.";
 RL Matrix Biol. 22:511-517(2003).
 DR EMBL; AY302252; AAP58398.1; --
 DR InterPro; IPR001173; Glyco_trans_2.
 DR Pfam; PF00535; Glycos_transf_1.
 SQ SEQUENCE 557 AA; 64060 MW; F81478B685BF7AD5 CRC64;
 Query Match 43.3%; Score 808; DB 2; Length 557;
 Best Local Similarity 53.9%; Pred. No. 3.2e-59;
 Matches 167; Conservative 40; Mismatches 75; Indels 28; Gaps 7;
 QY 17 SGLARRVLTAFAALLILGLMTWAYAAGVP--LASDRYGLLAFLGLYGAFLSAHLVAQSLFAY 75
 DB 7 TGL-RVLATCLFALLVLGGILVAVYVGYQFIHTEKHYSFLGYLGLHLLIQSLFAF 64
 QY 76 LEHRRVAAARGPLDARTARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVD 135
 DB 65 LEHRRMGGGRCPSGKST--VLCIAAYQEDPEYLRKLSRVRRLSYF--HLRIVMVD 119
 QY 136 GNRADLYWDMFREFVADDPATYVWGNHYQFWPEPAAGAVGAGAYREVEAEDPRLA 195
 DB 120 GNTTEDRYMDIFREVMGSEGTCCYIWDKNVHESEGGQEGERG----- 163
 QY 196 YEALVTRRCVCAQRWGGKREVMYTAFAKLGSDVDYVQVCDSDTRLDPMALLELVRLVLD 255
 DB 164 VOEMVKNFQYVVCIMQKMGKREVMYTAFAKLGSDVDYVQVCDSDTVLDPACTAEMLRILE 223
 QY 256 EDPVAGVGDVRIINPLDSWVSLSSRLRYVAFNVERACQSYFHCVSCISGPLESCPGP 315
 DB 224 EDPVAGVGDVQIILNKYDWSIFLSSVRYWMAFNVERACQSYFGCVQCISGFL-----GM 279

```

QY      316 REHAMPSFL 325
Db      280 YRNSLLQYFL 289

RESULT 10
HAS1_XENLA
ID      HAS1_XENLA      STANDARD;      PRT;      588 AA.
AC      P13563;
DT      01-JAN-1990 (Rel. 13, Created)
DT      01-JAN-1990 (Rel. 13, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE      (Hyaluronic acid synthase 1) (HA synthase 1) (XHASI) (DG42 protein).
GN      Name=HAS1;
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxID=8355;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=88311363; PubMed=3410156;
RA      Rosa F., Sargent T.D., Rebbert M.L., Michaels G.S., Jamrich M.,
RA      Grunz H., Jonas E., Winkler J.A., Dawid I.B.;
RT      "Accumulation and decay of DG42 gene products follow a gradient
RT      pattern during Xenopus embryogenesis.";
RL      Dev. Biol. 129:114-123(1988).
CC      !- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
CC      synthesis (By similarity). May play a role in signaling or pattern
CC      formation in embryonic development.
CC      !- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
CC      glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-
CC      glucuronosyl(1->3)](n) + 2n UDP.
CC      !- COFACTOR: Magnesium.
CC      !- PATHWAY: Hyaluronate synthesis.
CC      !- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC      !- TISSUE SPECIFICITY: Embryo.
CC      !- DEVELOPMENTAL STAGE: During gastrula and early neurula stages.
CC      !- SIMILARITY: Belongs to the nodC/HAS family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
EMBL; M22249; AAA49699.1; -;
DR      PIR; A43740; A43740.
DR      InterPro; IPR001173; Glyco trans 2.
KW      Glycosyltransferase; Multigene family; Transferase; Transmembrane.
FT      DOMAIN 1 28
FT      TRANSMEM 29 49
FT      DOMAIN 50 61
FT      TRANSMEM 62 82
FT      DOMAIN 83 411
FT      TRANSMEM 412 432
FT      DOMAIN 433 433
FT      TRANSMEM 434 454
FT      DOMAIN 455 456
FT      TRANSMEM 457 477
FT      DOMAIN 478 505
FT      TRANSMEM 506 526
FT      DOMAIN 527 543
FT      TRANSMEM 544 564
FT      DOMAIN 565 588
FT      TRANSMEM 588 AA; 68522 MW; 33DA3B8E331F4CE9 CRC64;
SQ      SEQUENCE 588 AA; 68522 MW; 33DA3B8E331F4CE9 CRC64;

Query Match      43.2%; Score 804.5; DB 1; Length 588;
Best Local Similarity      53.6%; Pred. No. 6.8e-59;
Matches 169; Conservative 44; Mismatches 79; Indels 21; Gaps 9;

```

```

Matches 156; Conservative 51; Mismatches 75; Indels 9; Gaps 3;
QY      22 RVLTIAPALLILGLMTWAYAAGVPLASDRYCLLAFGLYGAFLSAHLVAQSLFAYLEHRRV 81
Db      RIIYSEFVLLATITAAAYVAEFQVLKHEATLSGLYGLAMLLHMMQSLFAFLIEIRV 87
QY      82 AAAARGPLDAATARSVALTISAYOEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAED 141
Db      ---NKSELPCSPFKITVALTAGYOENFEYLKCLSECKYVYKPKDKLKIILVDGNTEDD 144
QY      142 LYVDMFREVFADEDPATYVMDGNHYHQFWEPAAGAGAGAGAYREVEA---EDPGRLAYEA 198
Db      ---AYMMEMKDFVHGSDVGTYYWKGNYHTVKKE---ETNKGSCPEVSKPLNDEGINMVEE 201
QY      199 LVTRRCVCVAQRWGGKREVMYTAFAKLGSDVDYVQVCDSTRLDPMALLRLVRLDSDP 258
Db      LVRNRCVCIMQWGGKREVMYTAFAIGTSVDYVQVCDSTKLDLATVEMVKVLESND 261
QY      259 RVGAVGGDVRILNPLDWSVFLSSRLRYWVAFNVERACQSYFHCVSCISGGL 309
Db      MYGAVGGDVRILNPLDYSFISFMSSRLRYWMAFNVERACQSYFDCVSCISGGL 312

RESULT 11
Q9DG40
ID      Q9DG40      PRELIMINARY;      PRT;      554 AA.
AC      Q9DG40;
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Hyaluronan synthase 3.
GN      Name=has3;
OS      Brachydanio rerio (Zebrafish) (Danio rerio).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Ostariphyssi; Cypriniformes;
OC      Cyprinidae; Danio.
OX      NCBI_TaxID=7955;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      PubMed=14729574;
RA      Bakkers J., Kramer C., Pothof J., Quaedvlieg N.E., Spaink H.P.,
RA      Hammerschmidt M.;
RT      "Has2 is required upstream of Rac1 to govern dorsal migration of
RT      lateral cells during zebrafish gastrulation.";
RL      Development 131:525-537(2004).
DR      EMBL; AF190743; AAG17181.1; -;
DR      ZFIN; ZDB-GENE-021118-1; has3.
DR      InterPro; IPR001173; Glyco trans 2.
DR      SEQUENCE 554 AA; 63077 MW; 26273CB459CE3E97 CRC64;

Query Match      43.1%; Score 803.5; DB 2; Length 554;
Best Local Similarity      54.0%; Pred. No. 7.7e-59;
Matches 169; Conservative 44; Mismatches 79; Indels 21; Gaps 9;
QY      18 GLABRV-LTAFALLILGLMTWAYAAGVPLASDRYCLLAFGLYGAFLSAHLVAQSLPAYL 76
Db      6 GTAVRIFITTLFAAVLVFAILLAYVTGQFIHTQHHLHFLGLYGAFLSLHLQLSLPAYL 65
QY      77 EHREVAAGARP-LDAATARSVALTISAYQBDPAYLRQCLASARALLYPRARLRVLMVVD 135
Db      ---RGPSRPQHLRRTVALCIAAYQEDPDYLRKCLSSR-ISP--GLKVVLVD 118
QY      136 GNRADLYMDMFEVFADEDPATYVMDGNHYHQFWEPAAGAGAGAGAYREVEADPRL 194
Db      GNRQEDAYMMDIFQEVNMGVGEQTGCVVMKGNHNSGDDGGGGGKGS-----VHAEEAARV 173
QY      195 AVEALVTRRCVCVAQRWGGKREVMYTAFAKLGSDVDYVQVCDSTRLDPMALLRLVRL 254
Db      ---RVVRSRCVSCIMQWGGKREVMYTAFAKLGSDVDYVQVCDSTVLDPACTEMLKIL 231
QY      255 DEDPRVAVGGDVRILNPLDWSVFLSSRLRYWVAFNVERACQSYFHCVSCISGPLESCPG 314
Db      EEDPDVGGVGGDVQILNKYDSWISFLSSRVYMAFNVERACQSYFGCVQCISGPL----G 287

```


QY 315 PREHAMPSFLAP 327
 Db 288 MYRNSLLQOFLP 300

RESULT 12

Q75R37
 ID Q75R37 PRELIMINARY; PRT; 553 AA.
 AC Q75R37;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hyaluronan synthase 3.
 GN Name=ehas3;
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kimura N., Konno Y., Yokoo M., Sato E.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB159675; BAD10881.1;
 DR InterPro; IPR001173; Glyco_trans_2.
 SQ SEQUENCE 553 AA; 63110 MW; C97EBE911723B44C CRC64;

Query Match 42.9%; Score 800.5; DB 2; Length 553;
 Best Local Similarity 54.1%; Pred. No. 1.4e-58;
 Matches 166; Conservative 38; Mismatches 84; Indels 19; Gaps 5;
 QY 21 RRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRR 80
 Db 10 RVVGTSLFALVGLGILAAVYTGQFIHTEKHYSFGLYGAILGLHLILQSLFAFLEHRR 69
 QY 81 VAAARG-PLDARTARSVALTISAYQEDPAYLRQCLASARALLYPPARLRVLMVVDGNR 139
 Db 70 MRRAGRPLKPLQSRVALCIAAYQEDPDYLRKCLRSQAQRIAFP--DLKVVVMVVDGNRQ 127
 QY 140 EDLYVMVDMFREV-ADDEDPATYVWDGNVHQPPEAAAGVAGAGAYREVEADPGRLAVEA 198
 Db 128 EDAYTMDLDFHEVLGGTEQAGFFVWRNSFNFEAGEGETEASLQEGMR-----VRN 176
 QY 199 LVTRRCVCVAQRWGGKREVMYTAFLKALGSDVDYVQVCDSDTRLDPMALLELRVLDDEP 258
 Db 177 VVRASTFCIMQKWKGGKREVMYTAFLKALGSDVDYVQVCDSDTVLDPACTFEMLRVLEEDP 236
 QY 259 RVGAVGGDVRIINPLDSWVSFLSSLRVYVAFNVERACQSYFHCVCISGPLESCPGPREH 318
 Db 237 QVGGVGGDVQILNKYDSWISFLSSVRYVWAFNVERACQSYFGCVQCISGPL-----GMVRN 292

QY 319 AMMPSPFL 325

Db 293 SLLQOFL 299

RESULT 13

HAS3 MOUSE
 ID HAS3 MOUSE STANDARD; PRT; 554 AA.
 AC O08650;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hyaluronan synthase 3 (PC 2.4.1.212) (Hyaluronate synthase 3)
 DE Hyaluronan synthase 3 (HA synthase 3).
 GN Name=Has3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvJ, and C57BL/6;

RX MEDLINE=97238820; PubMed=9083017; DOI=10.1074/jbc.272.14.8957;
 RA Spicer A.P., Olson J.S., McDonald J.A.;
 RT "Molecular cloning and characterization of a cDNA encoding the third
 RT putative mammalian hyaluronan synthase."; J. Biol. Chem. 272:8957-8961(1997).
 RL J. Biol. Chem. 272:8957-8961(1997).
 CC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA) synthesis.
 CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-glucuronosyl = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-glucuronosyl(1->3)](n) + 2n UDP.
 CC -!- COFACTOR: Magnesium.
 CC -!- PATHWAY: Hyaluronate synthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- DEVELOPMENTAL STAGE: Expressed at E17.5 day.
 CC -!- SIMILARITY: Belongs to the nodC/HAS family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U86408; AAC53128.1; -;
 DR MGD; MGI:109599; Has3.
 DR InterPro; IPR001173; Glyco_trans_2.
 KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
 FT DOMAIN 1 15 Cytoplasmic (Potential).
 FT TRANSMEM 16 36 Extracellular (Potential).
 FT DOMAIN 37 44 Extracellular (Potential).
 FT TRANSMEM 45 65 Cytoplasmic (Potential).
 FT DOMAIN 66 378 Cytoplasmic (Potential).
 FT TRANSMEM 379 399 Extracellular (Potential).
 FT DOMAIN 400 409 Extracellular (Potential).
 FT TRANSMEM 410 430 Cytoplasmic (Potential).
 FT DOMAIN 431 441 Cytoplasmic (Potential).
 FT TRANSMEM 442 462 Extracellular (Potential).
 FT DOMAIN 463 474 Extracellular (Potential).
 FT TRANSMEM 475 495 Cytoplasmic (Potential).
 FT DOMAIN 496 516 Cytoplasmic (Potential).
 FT TRANSMEM 517 537 Extracellular (Potential).
 FT DOMAIN 538 554 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 463 463
 SQ SEQUENCE 554 AA; 63338 MW; 88296DCB465CE3EC CRC64;
 Query Match 42.9%; Score 800; DB 1; Length 554;
 Best Local Similarity 53.7%; Pred. No. 1.5e-58;
 Matches 166; Conservative 41; Mismatches 80; Indels 22; Gaps 6;
 QY 21 RRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRR 80
 Db 10 RVVGTSLFALVGLGILAAVYTGQFIHTEKHYSFGLYGAILGLHLILQSLFAFLEHRR 69
 QY 81 VAAARGPLD---AATARSVALTISAYQEDPAYLRQCLASARALLYPPARLRVLMVVDGN 137
 Db 70 MRRAGR-FLKLCRSQSRVALCIAAYQEDPEYLRKCLRSQAQRIAFP--NLKVVVMVVDGN 126
 QY 138 RAEDLYVMVDMFREV-ADDEDPATYVWDGNVHQPPEAAAGVAGAGAYREVEADPGRLAV 196
 Db 127 RQEDTYMDLDFHEVLGGTEQAGFFVWRNSFNFEAGEGETEASLQEGMR-----V 175
 QY 197 EALVTRRCVCVAQRWGGKREVMYTAFLKALGSDVDYVQVCDSDTRLDPMALLELRVLD 256
 Db 176 RAVVWASTFCIMQKWKGGKREVMYTAFLKALGSDVDYVQVCDSDTVLDPACTFEMLRVLE 235
 QY 257 DPRVAVGGDVRIINPLDSWVSFLSSLRVYVAFNVERACQSYFHCVCISGPLESCPGPR 316
 Db 236 DPQVGGVGGDVQILNKYDSWISFLSSVRYVWAFNVERACQSYFGCVQCISGPL-----GMV 291
 QY 317 EHAMPSPFL 325
 Db 292 RNSLLQOFL 300


```

QY 81 VAAAARGPLD---AATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRLMVVDGN 137
Db : | | | : | | | | | | | | | | | | | | | | | | | | | | | |
70 MRRAGR-PLKLRCSQRRRSVALCIAAYQEDPEYLRKCLRSQRIAPP--NLKVVVVVDGN 126
QY 138 RAEDLYMVDMPREVF-ADEDPATYVWDGNYHQWEPAPAAAGAVGAGAYREVEAEDPRLAV 196
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
127 ROEDAYMLDIFHEVLGGTEQAGFFVVRGNFHEAGEGETEASLQEGMER-----V 175
QY 197 EALVRTRRCVCAQRWGGKREVMYTAFAKALGDSVDYVOVCDSDTRLDPMALLELVRVLDE 256
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
176 RAVVWASTFSCINQKXWGGKREVMYTAFAKALGNSVDYIQVCDSDTVLDPACTIEMLRVLEE 235
QY 257 DPRVAVGGDVRIILNPLDSWISFLSLRYWVAFNVERACQSYFHCVSCISGPLESCGPR 316
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
236 DPQVGGVGGDVQILNKYDSWISFLSSRYWMAFNVERACQSYFGCVQCISGFL----GMY 291
QY 317 EHAMPSPFL 325
Db : | | | | |
292 RNSLIQQFL 300

```

Search completed: March 11, 2005, 14:21:58
Job time : 96.0909 secs

THIS PAGE BLANK (US10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:11:23 ; Search time 26.25 Seconds
(without alignments)
1023.759 Million cell updates/sec

Title: US-10-672-399-6
Perfect score: 1864
Sequence: 1 MRQDAPKPTPAARRCSGLA.....ALVLPVAEPADTLVQVLP 360

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1505	86.1	578	US-08-635-552A-2	Sequence 2, Appli
2	1597	85.7	589	US-09-949-016-11530	Sequence 11530, A
3	1512	81.1	583	US-08-675-499A-3	Sequence 3, Appli
4	1512	81.1	583	US-08-812-008-3	Sequence 3, Appli
5	1429	76.7	543	US-09-155-768-4	Sequence 4, Appli
6	804.5	43.2	587	US-08-635-552A-3	Sequence 3, Appli
7	800	42.9	554	US-08-812-008-32	Sequence 32, Appl
8	790.5	42.4	553	US-09-949-016-9599	Sequence 9599, Ap
9	788	42.3	587	US-08-675-499A-4	Sequence 4, Appli
10	788	42.3	587	US-08-812-008-4	Sequence 4, Appli
11	783.5	42.0	552	US-08-675-499A-2	Sequence 2, Appli
12	783.5	42.0	552	US-08-812-008-2	Sequence 2, Appli
13	783.5	42.0	552	US-09-949-016-6608	Sequence 6608, Ap
14	463.5	24.9	241	US-08-865-273-2	Sequence 2, Appli
15	463.5	24.9	241	US-09-385-174-2	Sequence 2, Appli
16	310	16.6	568	US-09-469-200E-10	Sequence 10, Appl
17	217	11.6	43	US-08-675-499A-9	Sequence 9, Appli
18	217	11.6	43	US-08-812-008-9	Sequence 9, Appli
19	204.5	11.0	426	US-08-675-499A-6	Sequence 6, Appli
20	204.5	11.0	426	US-08-812-008-6	Sequence 6, Appli
21	177	9.5	43	US-08-812-008-35	Sequence 35, Appl
22	174	9.3	417	US-09-469-200E-2	Sequence 2, Appli
23	173	9.3	43	US-08-675-499A-7	Sequence 7, Appli
24	173	9.3	43	US-08-812-008-7	Sequence 7, Appli
25	163	8.7	43	US-08-675-499A-10	Sequence 10, Appl
26	163	8.7	43	US-08-812-008-10	Sequence 10, Appl
27	151.5	8.1	419	US-08-270-581-2	Sequence 2, Appli

28	151.5	8.1	419	4	US-09-146-893-2	Sequence 2, Appli
29	151.5	8.1	419	4	US-08-675-499A-5	Sequence 5, Appli
30	151.5	8.1	419	4	US-08-812-008-5	Sequence 5, Appli
31	149.5	8.0	395	4	US-08-635-552A-4	Sequence 4, Appli
32	139	7.5	1093	4	US-09-248-796A-17108	Sequence 17108, A
33	135.5	7.3	393	4	US-09-902-540-11514	Sequence 11514, A
34	117.5	6.3	403	4	US-09-902-540-11529	Sequence 11529, A
35	110.5	5.9	904	4	US-09-252-991A-19257	Sequence 19257, A
36	104.5	5.6	419	4	US-09-328-352-7089	Sequence 7089, Ap
37	104.5	5.6	3519	3	US-09-428-517-4	Sequence 4, Appli
38	104	5.6	738	4	US-09-543-681A-7528	Sequence 7528, Ap
39	100.5	5.4	4150	3	US-09-428-517-2	Sequence 2, Appli
40	99.5	5.3	418	3	US-09-134-001C-4051	Sequence 4051, Ap
41	99.5	5.3	3562	4	US-09-679-279-14	Sequence 14, Appl
42	95.5	5.1	485	4	US-09-902-540-11019	Sequence 11019, A
43	95.5	5.1	1380	4	US-09-252-991A-25722	Sequence 25722, A
44	95	5.1	254	4	US-09-248-796A-16421	Sequence 16421, A
45	94	5.0	831	4	US-09-252-991A-24901	Sequence 24901, A

ALIGNMENTS

RESULT 1
US-08-635-552A-2
; Sequence 2, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millicia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-635-552A-2

Query Match 86.1%; Score 1605; DB 4; Length 578;
Best Local Similarity 96.0%; Pred. No. 2.8e-179;
Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;
QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLADRYGLLAFLYLG 60
Db 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLADRYGLLAFLYLG 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLFQCLASARA 120
|||||

Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Qy 121 LLYPRARLVLMVVDGNRAEDLYMDMPREVFADEDPATYVWDGNYHQWPEPAAAGAYGA 180
Db 121 LLYPRARLVLMVVDGNRAEDLYMDMPREVFADEDPATYVWDGNYHQWPEPAAAGAYGA 180
Qy 181 GAYREVEADPGRLAVEALVRTRRCVCAORWGGKREVMYTAFAKALGDSVDVQVCDSDT 240
Db 181 GAYREVEADPGRLAVEALVRTRRCVCAORWGGKREVMYTAFAKALGDSVDVQVCDSDT 240
Qy 241 RLDPMALLELRVLDDEDPVGAAGDVRILNPLDSWVSFLSLRYWAFNVERACQSYFH 300
Db 241 RLDPMALLELRVLDDEDPVGAAGDVRILNPLDSWVSFLSLRYWAFNVERACQSYFH 300
Qy 301 CVSCISGPLESCEPGPREHAMPSFL 325
Db 301 CVSCISGGLP-----GLYRNLLQOFL 321

RESULT 2
US-09-949-016-11530
; Sequence 11530, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11530
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11530

Query Match 85.7%; Score 1597; DB 4; Length 589;
Best Local Similarity 95.7%; Pred. No. 2.5e-178;
Matches 311; Conservative 3; Mismatches 7; Indels 4; Gaps 1;
Qy 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYG 60
Db 12 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYG 71
Qy 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db 72 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 131
Qy 121 LLYPRARLVLMVVDGNRAEDLYMDMPREVFADEDPATYVWDGNYHQWPEPAAAGAYGA 180
Db 132 LLYPRARLVLMVVDGNRAEDLYMDMPREVFADEDPATYVWDGNYHQWPEPAAAGAYGA 191
Qy 181 GAYREVEADPGRLAVEALVRTRRCVCAORWGGKREVMYTAFAKALGDSVDVQVCDSDT 240
Db 192 GAYREVEADPGRLAVEALVRTRRCVCAORWGGKREVMYTAFAKALGDSVDVQVCDSDT 251
Qy 241 RLDPMALLELRVLDDEDPVGAAGDVRILNPLDSWVSFLSLRYWAFNVERACQSYFH 300
Db 252 RLDPMALLELRVLDDEDPVGAAGDVRILNPLDSWVSFLSLRYWAFNVERACQSYFH 311
Qy 301 CVSCISGPLESCEPGPREHAMPSFL 325
Db 312 CVSCISGGLP-----GLYRNLLQOFL 332

RESULT 3
US-08-675-499A-3
; Sequence 3, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-499A-3

Query Match 81.1%; Score 1512; DB 4; Length 583;
Best Local Similarity 90.6%; Pred. No. 2.4e-168;
Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3;
Qy 3 QODAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAF 62
Db 2 RQDMPKPSAARCCSGLARRALTIIFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAF 61
Qy 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 117
Db 62 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 121
Qy 118 ARALLYPRARLVLMVVDGNRAEDLYMDMPREVFADEDPATYVWDGNYHQWPEPAAAG 176
Db 122 ARALLYPHTLRVLMVVDGNRAEDLYMDMPREVFADEDPATYVWDGNYHQWPEPAAAG 181
Qy 177 AVGAGAYREVEADPGRLAVEALVRTRRCVCAORWGGKREVMYTAFAKALGDSVDVQVQC 236
Db 182 AVGAGAYREVEADPGRLAVEALVRTRRCVCAORWGGKREVMYTAFAKALGDSVDVQVQC 241
Qy 237 DSDTRLDPMALLELRVLDDEDPVGAAGDVRILNPLDSWVSFLSLRYWAFNVERACQ 296
Db 242 DSDTRLDPMALLELRVLDDEDPVGAAGDVRILNPLDSWVSFLSLRYWAFNVERACQ 301
Qy 297 SYFHCVSCISGPLESCEPGPREHAMPSFL 325

Db 302 SYFHCVCISGSLP-----GLYRNNLLQQL 326

RESULT 4

US-08-812-008-3

; Sequence 3, Application US/08812008

; Patent No. 6602693

; GENERAL INFORMATION:

; APPLICANT: McDonald, J. A.

; APPLICANT: Spicer, A. P.

; APPLICANT: Augustine, M. L.

; TITLE OF INVENTION: GENE ENCODING HYALURONAN

; TITLE OF INVENTION: SYNTHASE

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Schwegman, Lundberg, Moessner & Kluth, P.A.

; STREET: P.O. Box 2938

; CITY: Minneapolis

; STATE: MN

; COUNTRY: U.S.A

; ZIP: 55402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/812,008

; FILING DATE: 05-MAR-1997

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/675,499

; FILING DATE: 03-JUL-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Embretson, Janet E

; REGISTRATION NUMBER: 39,665

; REFERENCE/DOCKET NUMBER: 150.183US1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 612-373-6900

; TELEFAX: 612-339-3061

; TELEX:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 583 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-812-008-3

Query Match 81.1%; Score 1512; DB 4; Length 583;

Best Local Similarity 90.6%; Pred. No. 2.4e-168;

Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3;

Qy 3 QODAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFLYGAF 62

Db 2 RQDMKPESEARRCCSGLARRALTIFALLILGLMTWYAAGVPLASDRYGLLAFLYGAF 61

Qy 63 LSAHLVAQSLFAYLEHRRVAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCLAS 117

Db 62 LSAHLVAQSLFAYLEHRRVAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCLTS 121

Qy 118 ARALLYPRLRLVLMVVDGNRAEDLYVDMPREVEADEDPATYVWDGNVHQPWBPAAAG 176

Db 122 ARALLYPHTRRLVLMVVDGNRAEDLYVDMPREVEADEDPATYVWDGNVHQPWBPAAAG 181

Qy 177 AVGAGAYREVEADEPGRLAVALVTRRCVCVAQRWGKREVMYTAFAKALGSDVDYVQVC 236

Db 182 AVGAGAYREVEADEPGRLAVALVTRRCVCVAQRWGKREVMYTAFAKALGSDVDYVQVC 241

Qy 237 DSDTRLDPMALLELRVLDDEPRVAGVGDVRLNPLDSWSFSLSLRYWAFNVERACQ 296

Db 242 DSDTRLDPMALLELRVLDDEPRVAGVGDVRLNPLDSWSFSLSLRYWAFNVERACQ 301

Qy 297 SYFHCVCISGPLESCPGPREHAMMPSFL 325

Db 302 SYFHCVCISGPL-----GLYRNNLLQQL 326

RESULT 5

US-09-155-768-4

; Sequence 4, Application US/09155768A

; Patent No. 6162908

; GENERAL INFORMATION:

; APPLICANT: SEIKAGAKU CORPORATION

; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF

; FILE REFERENCE: HUMAN-ORIGIN AND DNA ENCODING THE SAME

; FILE REFERENCE: TOYAM37.001APC

; CURRENT APPLICATION NUMBER: US/09/155,768A

; CURRENT FILING DATE: 1998-10-05

; EARLIER APPLICATION NUMBER: JP 8-084326

; EARLIER FILING DATE: 1996-04-05

; EARLIER APPLICATION NUMBER: JP 8-109663

; EARLIER FILING DATE: 1996-04-30

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 4

; LENGTH: 543

; TYPE: PRT

; ORGANISM: HUMAN

US-09-155-768-4

Query Match 76.7%; Score 1429; DB 3; Length 543;

Best Local Similarity 95.2%; Pred. No. 1.2e-158;

Matches 276; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

Qy 36 MTWYAAGVPLASDRYGLLAFLYCAFLSAHLVAQSLFAYLEHRRVAAARGLDAATAR 95

Db 1 MTWYAAGVPLASDRYGLLAFLYCAFLSAHLVAQSLFAYLEHRRVAAARGLDAATAR 60

Qy 96 SVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAEDLYVDMPREVFADE 155

Db 61 SVALTISAYQEDPAYLRQCLASARALLYPRARVRVLMVVDGNRAEDLYVDMPREVFADE 120

Qy 156 DPATYVWDGNVHQPWBPAAAGAGAGAYREVEADEPGRLAVALVTRRCVCVAQRWGK 215

Db 121 DPATYVWDGNVHQPWBPAAAGAGAGAYREVEADEPGRLAVALVTRRCVCVAQRWGK 180

Qy 216 REVMTAFKALGSDVDYVQVCSDTRLDPMALLELRVLDDEPRVAGVGDVRLNPLDS 275

Db 181 REVMTAFKALGSDVDYVQVCSDTRLDPMALLELRVLDDEPRVAGVGDVRLNPLDS 240

Qy 276 WVSFLSLRYWAFNVERACQSYFHCVCISGPLESCPGPREHAMMPSFL 325

Db 241 WVSFLSLRYWAFNVERACQSYFHCVCISGPL-----GLYRNNLLQQL 286

RESULT 6

US-08-635-552A-3

; Sequence 3, Application US/08635552A

; Patent No. 6423514

; GENERAL INFORMATION:

; APPLICANT: Briskin, Michael J.

; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic

; TITLE OF INVENTION: Acids and Uses Thereof

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/635,552A
FILING DATE: 22-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-07
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 587 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-635-552A-3

Query Match 43.2%; Score 804.5; DB 4; Length 587;
Best Local Similarity 53.6%; Pred. No. 3.5e-85;
Matches 156; Conservative 51; Mismatches 75; Indels 9; Gaps 3;
QY 22 RVLTIAFALLILGLMTWAVAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLHRRV 81
DB RIYYSGVLLATIAAYVAFQVLKHEALFSLGLYGLAWLLHMMQSLFAFLHRRV 87
QY 82 AAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAED 141
DB 88 ---NKSELPCKFKTVALTIAGYQENPEYLKLCLECKVVKYKPKDKLIIILVIDGNTDD 144
QY 142 LYMDMFREVFADEDPATYVWDGNTHQWPEPAAAGAVGAYREVEA---BDPGRLAVEA 198
DB 145 AYMEMFMDVFGEDVGTVMKGNHTVKKPE---ETNKGSCPEVSKPLNDEGINMVEE 201
QY 199 LVRTRCVCVQQRWKGKREVMYTAFAKALGDSVDYVQVCDSDTRDLPMLLELVRVLDEP 258
DB 202 LVRNKRVCIMQWQKGKREVMYTAFAIGTSVDYVQVCDSDTKDELATVEMVKVLESND 261
QY 259 RVGAVGGDVRIINPLDSDWVSFLSSLRVYVAFNVERACQSYFHCVSCISGSL 309
DB 262 MYGAVGGDVRIINPLDSDWVSFLSSLRVYVAFNVERACQSYFHCVSCISGSL 312

RESULT 7
US-08-812-008-32
Sequence 32, Application US/08812008
Patent No. 6602693
GENERAL INFORMATION:
APPLICANT: McDonald, J. A.
APPLICANT: Spicer, A. P.
APPLICANT: Augustine, M. L.
TITLE OF INVENTION: GENE ENCODING HYALURONAN
TITLE OF INVENTION: SYNTHASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: U.S.A.
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,008
FILING DATE: 05-MAR-1997

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/675,499
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Embretson, Janet E.
REGISTRATION NUMBER: 39,665
REFERENCE/DOCKET NUMBER: 150.183US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 554 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-812-008-32

Query Match 42.9%; Score 800; DB 4; Length 554;
Best Local Similarity 53.7%; Pred. No. 1.1e-84;
Matches 166; Conservative 41; Mismatches 80; Indels 22; Gaps 6;
QY 21 RRVLTIAFALLILGLMTWAVAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLHRR 80
DB 10 RVGTSFALVVLGGILAAVYTGQFTHTEKHYSFGLYGAILGLHLTIQSLFAPLEHRR 69
QY 81 VAAARGPLD---AATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGN 137
DB 70 MRAGR-PLKLHCSQRSVALCIAAQEDPEYLKCLRSQRIAPP--NLKVVVVVDGN 126
QY 138 RAEDLYMVDMPREVF-ADDPATYVWDGNTHQWPEPAAAGAVGAYREVEAEDPGRIV 196
DB 127 RQEDTYMLDIFHEVLGGTEQAGFVWRSNFHEAGEGETEASLQGMER-----V 175
QY 197 EALVTRCVCVQQRWKGKREVMYTAFAKALGDSVDYVQVCDSDTRDLPMLLELVRVLDE 256
DB 176 RAVVWASTFCIMQWQKGKREVMYTAFAKALGNSVDYIQVCDSDVLDPACTIEMLRVLEE 235
QY 257 DPRVAVGGDVRIINPLDSDWVSFLSSLRVYVAFNVERACQSYFHCVSCISGSLSCPGPR 316
DB 236 DPQVGGVGGDVQILNKYDSWISFLSSRVYVWMAFNVERACQSYFCVCQISGSL----GMY 291
QY 317 EHMMPSFL 325
DB 292 RNSLLQQL 300

RESULT 8
US-09-949-016-9599
Sequence 9599, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9599
LENGTH: 553
TYPE: PRT


```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-812-008-4

Query Match 42.3%; Score 788; DB 4; Length 587;
Best Local Similarity 53.3%; Pred. No. 3.1e-83;
Matches 155; Conservative 51; Mismatches 75; Indels 10; Gaps 4;

Qy 22 RVLTIAPALLILGLMTWYAAAGVP-LASDRYGLLAGFLYAGFLSAHLVAQSLFAYLEHRRV 81
Db 28 RIIVYFSGWLLATITAAAYAEFQVLKHEALFSLGLYLGLAMLLHMMQSLFAFLFIRRV 87
Qy 82 AAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRLVLMVVDGNRAED 141
Db 88 ---NKSELPFCFKTVALTIAGYQENPEYLLKLESCKYVYKPKGLKIIILVIDGNTEDD 144
Qy 142 LYMDMPREVFADEDPATYVWDGNHQPWEPAAGAVGAGAYREVEA---EDPGRLAVEA 198
Db 145 AYMEMPFDVPHGEDVGYVWKNYHTYKKEP---ETNKGSCPEVSKPLNEDEGINNVVEE 201
Qy 199 LVRTRRCVCVAQRMGKREVMYTAFAKALGDSVYVQVCDSTRLDPMALLELRLVRLDEDP 258
Db 202 LVRNKRVCVIMQW-GKREVMYTAFAIGTSVYVQVCDSTKDLDELATVEMVVKLESND 260
Qy 259 RVGAVGGDVRILNPLDSWVSLSLRYWVAFNVERACOSYFHCVCISGGL 309
Db 261 MYGAVGGDVRILNPNYDSFISPMSSSLRYWMAFNVERACOSYFDCVCSISGGL 311

RESULT 11
US-08-675-499A-2
; Sequence 2, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-499A-2

Query Match 42.0%; Score 783.5; DB 4; Length 552;
Best Local Similarity 51.3%; Pred. No. 9.4e-83;
Matches 157; Conservative 52; Mismatches 76; Indels 21; Gaps 7;

Qy 21 RRVLTIAFALLILGLMTWYAAAGVP-LASDRYGLLAGFLYAGFLSAHLVAQSLFAYLEHR 79
Db 11 RIITITLFGVSLGITAAYIVGQFOTDNY-YFSGLYAGFLASHLIQSLFAFLFHR 69
Qy 80 RVAAGRGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRLVLMVVDGNRA 139
Db 70 KMKSLSTFPI--KLNTKVALCIAAYQEDPYLRKCLQSVKELTYP--GIKVMVWDGNSD 125
Qy 140 EDLYMDMPREVFADEDPATYVWDGNHQPWEPAAGAVGAGAYREVEAEDPGRLAVEAL 199
Db 126 DDLYMDIFSEVIGRDKSATYVWKNPFHEK-----GPGETEESHKSSQH--VTQL 174
Qy 200 VTRTRCVCVAQRMGKREVMYTAFAKALGDSVYVQVCDSTRLDPMALLELRLVRLDEDP 259
Db 175 VLSNKSICIMQKMGKREVMYTAFAKALGDSVYVQVCDSTMLDPASSVEMVKVLEEDPM 234
Qy 260 VGAVGGDVRILNPLDSWVSLSLRYWVAFNVERACOSYFHCVCISGGLSCPGPREHA 319
Db 235 VGVGGDVRILNPKYDSWISFLSSVRYWMAFNIERACOSYFGCVQICISGPL-----GMYRNS 290
Qy 320 MPPSFL 325
Db 291 LLHEFV 296

RESULT 12
US-08-812-008-2
; Sequence 2, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 552 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-812-008-2

Query Match 42.0%; Score 783.5; DB 4; Length 552;
Best Local Similarity 51.3%; Pred. No. 9.4e-83;
Matches 157; Conservative 52; Mismatches 76; Indels 21; Gaps 7;

QY 21 RRVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFLGAGFLSAHLVAQSLFAYLEHR 79
DB 11 RIIGTTLFGVSLGLGITAAYIVGYQFIQTDNY-YFSFGLYGAFLASHLIQSLFAFLBHR 69
QY 80 RVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRA 139
DB 70 KMKKSLETPI--KLNKTVALCIAAQEDDPYLRKCLQSVKRLTYP--GIKVVNVIDGNSD 125
QY 140 EDLYMVMDFREFADEDPATYYVMDGNYHQPWEPAAGAGVAGAYREVEAEDPGRILAVAL 199
DB 126 DDLYMMDIFSEVGMGRDYSATYIWKNNFHEK-----GPGTEESHKESQH--VTQL 174
QY 200 VTRRCVCVQAORWGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMALLELVRLVLEDPDR 259
DB 175 VLSNKSICIMQKWKGRVMTAFRALGRSDYVQVCDSDTMDLPASSVEMVKVLEEDPM 234
QY 260 VGAVGDDVRIINPLDSWVSFLSLRYWVAFNVERACQSYFHCVCISGPLESCPGPREHA 319
DB 235 VGVGDDVQIILNKYDSWISFLSVRYWMAFNIERACQSYFGVCQCSISGPL-----GMVRS 290
QY 320 MNPSPFL 325
DB 291 NSLLHEFV 296

RESULT 13

US-09-949-016-6608
Sequence 6608, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6608
LENGTH: 552
TYPE: PRT
ORGANISM: Human
US-09-949-016-6608

Query Match 42.0%; Score 783.5; DB 4; Length 552;
Best Local Similarity 51.0%; Pred. No. 9.4e-83;
Matches 157; Conservative 53; Mismatches 77; Indels 21; Gaps 7;

QY 19 LARRVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFLGAGFLSAHLVAQSLFAYLE 77
DB 9 ILRIIGTTLFGVSLGLGITAAYIVGYQFIQTDNY-YFSFGLYGAFLASHLIQSLFAFL 67
QY 78 HRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGN 137
DB 68 HRKMKKSLTPTI--KLNKTVALCIAAQEDDPYLRKCLQSVKRLTYP--GIKVVNVIDGN 123

QY 138 RAEDLYMVMDFREFADEDPATYYVMDGNYHQPWEPAAGAGVAGAYREVEAEDPGRILAVE 197
DB 124 SEDDLYMMDIFSEVGMGRDYSATYIWKNNFHEK-----GPGTEESHKESQH--VT 172
QY 198 ALVTRRCVCVQAORWGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMALLELVRLVLE 257
DB 173 QLVLSNKSICIMQKWKGRVMTAFRALGRSDYVQVCDSDTMDLPASSVEMVKVLEBED 232
QY 258 PRVGAAGDDVRIINPLDSWVSFLSLRYWVAFNVERACQSYFHCVCISGPLESCPGPRE 317
DB 233 PMVGDDVQIILNKYDSWISFLSVRYWMAFNIERACQSYFGVCQCSISGPL-----GMV 288
QY 318 HAMMPSFL 325
DB 289 NSLLHEFV 296

RESULT 14

US-08-865-273-2
Sequence 2, Application US/08865273
Patent No. 5994100
GENERAL INFORMATION:
APPLICANT: ZHU, YUAN
APPLICANT: NAMBI, PONNAL
APPLICANT: PULLEN, MARK A
TITLE OF INVENTION: NOVEL HAS2 SPLICING VARIANT
TITLE OF INVENTION: HOBEFC11: A TARGET IN CHRONIC RENAL FAILURE,
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESSES:
ADDRESS: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fast-Seq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,273
FILING DATE: 29-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70053
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 241 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-865-273-2

Query Match 24.9%; Score 463.5; DB 2; Length 241;
Best Local Similarity 45.4%; Pred. No. 1e-45;
Matches 99; Conservative 39; Mismatches 63; Indels 17; Gaps 6;

QY 19 LARRVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFLGAGFLSAHLVAQSLFAYLE 77
DB 9 ILRIIGTTLFGVSLGLGITAAYIVGYQFIQTDNY-YFSFGLYGAFLASHLIQSLFAFL 67
QY 78 HRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGN 137

Db 68 HRKMKSLPTI--KLNKTVLALCIAAYQEDPDYLRKCLQSVKRLTYP--GIKVVWVIDGN 123
QY 138 RAEDLYMVDMPREVFADDPATYVWDGNYHPWEPAAGAGAGAYREVEAEDPGLAVE 197
Db 124 SEDDLYMMDIFSEVWGRDKSATHIWKNNFHEK-----GPGETDESHKSSQH--VT 172
QY 198 ALVTRTRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQV 235
Db 173 QLVLSNKSICIMQKWKGGKREVMYTAFAKALGDSVDYVQV 210

RESULT 15

US-09-385-174-2
; Sequence 2, Application US/09385174
; Patent No. 6350446
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN
; NAMBI, PONNAL
; PULLEN, MARK A
; TITLE OF INVENTION: NOVEL HAS2 SPLICING VARIANT
; HOEFCL1: A TARGET IN CHRONIC RENAL FAILURE,
; INFLAMMATORY DISEASES AND MYOCARDIAL ISCHEMIA
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/385,174
; FILING DATE: 30-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/865,273
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-385-174-2

Query Match 24.9%; Score 463.5; DB 3; Length 241;
Best Local Similarity 45.4%; Pred. No. 1e-45;
Matches 99; Conservative 39; Mismatches 63; Indels 17; Gaps 6;
QY 19 LARRVLTIAFALLIIGLMTWAYAGVP-LASDRYGLLAFLGYLAFLSAHLVAQSLFAYLE 77
Db 9 ILRIIGTTLFGVSLLLGITAIVYQYQITDNY-YFSFGLYGAFSLASHLIQSLFAPLE 67
QY 78 HRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCCLASARALLYPRARLRVLMVDGN 137
Db 68 HRKMKSLPTI--KLNKTVLALCIAAYQEDPDYLRKCLQSVKRLTYP--GIKVVWVIDGN 123

QY 138 RAEDLYMVDMPREVFADDPATYVWDGNYHPWEPAAGAGAGAYREVEAEDPGLAVE 197
Db 124 SEDDLYMMDIFSEVWGRDKSATHIWKNNFHEK-----GPGETDESHKSSQH--VT 172
QY 198 ALVTRTRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQV 235
Db 173 QLVLSNKSICIMQKWKGGKREVMYTAFAKALGDSVDYVQV 210

Search completed: March 11, 2005, 14:24:37
Job time : 31.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:22:15 ; Search time 80.7955 Seconds
(without alignments)
1469.688 Million cell updates/sec

Title: US-10-672-399-6
Perfect score: 1864
Sequence: 1 MRQDAPKPTPAARRCSGLA.....ALVPAVAEPADTLVQVVLP 360

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844858 residues
Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1864	100.0	360	16	US-10-672-399-6
2	1612	86.5	376	16	US-10-672-399-4
3	1605	86.1	578	13	US-10-042-523-2
4	1605	86.1	578	16	US-10-672-399-2
5	1512	81.1	583	10	US-09-902-939-4
6	1298	69.6	320	16	US-10-672-399-8
7	804.5	43.2	587	13	US-10-042-523-3
8	804.5	43.2	588	15	US-10-309-560-10
9	800	42.9	554	10	US-09-902-939-6
10	790.5	42.4	553	15	US-10-295-027-370
11	790.5	42.4	553	15	US-10-188-832-137
12	783.5	42.0	552	10	US-09-902-939-5
13	783.5	42.0	552	14	US-10-262-526-2

14	783.5	42.0	552	14	US-10-262-526-4	Sequence 4, Appli
15	310	16.6	567	9	US-09-879-959-7	Sequence 7, Appli
16	310	16.6	567	14	US-10-172-527-7	Sequence 7, Appli
17	310	16.6	567	15	US-10-309-560-12	Sequence 12, Appli
18	310	16.6	568	14	US-10-011-768B-10	Sequence 10, Appli
19	310	16.6	568	14	US-10-011-771B-10	Sequence 10, Appli
20	174	9.3	417	9	US-09-879-959-2	Sequence 2, Appli
21	174	9.3	417	14	US-10-011-768B-2	Sequence 2, Appli
22	174	9.3	417	14	US-10-011-771B-2	Sequence 2, Appli
23	174	9.3	417	14	US-10-172-527-2	Sequence 2, Appli
24	174	9.3	417	14	US-10-326-185-2	Sequence 2, Appli
25	174	9.3	417	15	US-10-309-560-2	Sequence 2, Appli
26	174	9.3	417	15	US-10-309-560-16	Sequence 16, Appli
27	174	9.3	417	15	US-10-309-560-17	Sequence 17, Appli
28	174	9.3	417	15	US-10-309-560-18	Sequence 18, Appli
29	174	9.3	417	15	US-10-309-560-22	Sequence 22, Appli
30	174	9.3	417	15	US-10-309-560-23	Sequence 23, Appli
31	174	9.3	417	15	US-10-309-560-24	Sequence 24, Appli
32	174	9.3	417	15	US-10-309-560-25	Sequence 25, Appli
33	170	9.1	417	15	US-10-309-560-15	Sequence 15, Appli
34	170	9.1	417	15	US-10-309-560-19	Sequence 19, Appli
35	170	9.1	417	15	US-10-309-560-20	Sequence 20, Appli
36	170	9.1	417	15	US-10-309-560-21	Sequence 21, Appli
37	170	9.1	417	15	US-10-309-560-26	Sequence 26, Appli
38	170	9.1	417	15	US-10-309-560-27	Sequence 27, Appli
39	170	9.1	417	15	US-10-309-560-28	Sequence 28, Appli
40	170	9.1	417	15	US-10-309-560-29	Sequence 29, Appli
41	159	8.5	1219	15	US-10-389-493-3212	Sequence 3212, Ap
42	156.5	8.4	1165	9	US-09-801-368-76	Sequence 76, Appli
43	156.5	8.4	1165	15	US-10-369-493-1432	Sequence 1432, Ap
44	156	8.4	1160	15	US-10-369-493-12860	Sequence 12860, A
45	152.5	8.2	419	15	US-10-309-560-30	Sequence 30, Appli

ALIGNMENTS

RESULT 1
US-10-672-399-6
; Sequence 6, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-6

Query Match 100.0%; Score 1864; DB 16; Length 360;
Best Local Similarity 100.0%; Pred. No. 2.8e-178;
Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAGVPLASDRYGLLAFLYLG	60
Db	1	MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAGVPLASDRYGLLAFLYLG	60
Qy	61	AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCCLASARA	120
Db	61	AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCCLASARA	120
Qy	121	LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQPWEPAAAGAVGA	180
Db	121	LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQPWEPAAAGAVGA	180

181	Qy	GAYREVEAEDPGLRAVEALVTRTRCVCVAQRMWGKREVMYTAFAKALGDSVDYVYQVCDSDT	240
181	Db	GAYREVEAEDPGLRAVEALVTRTRCVCVAQRMWGKREVMYTAFAKALGDSVDYVYQVCDSDT	240
241	Qy	RLDPWALLELVRLVLEDDPRVAGVGDDVRLNPLDSWVSFLSSRLRYWVAFNVERACQSYFH	300
241	Db	RLDPWALLELVRLVLEDDPRVAGVGDDVRLNPLDSWVSFLSSRLRYWVAFNVERACQSYFH	300
301	Qy	CVSCTSGPLESCEPGREHAMMPSFLAPVOVHLQVPLLIRDALVLPAAVEADPTLVQVWLP	360
301	Db	CVSCTSGPLESCEPGREHAMMPSFLAPVOVHLQVPLLIRDALVLPAAVEADPTLVQVWLP	360

RESULT 2

US-10-672-399-4
; Sequence 4, Application US/10672399
; Publication No. US20050003368A1

; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics

: FILE REFERENCE: A89403303
: CURRENT APPLICATION NUMBER: US/10/672,399

; CURRENT AFFILIATION NUMBER: 03/1
 ; CURRENT FILING DATE: 2003-09-25

; PRIORITY APPLICATION NUMBER: US 60/472,401

; PRIOR FILING DATE: 2003-05-22

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 4

; LENGTH: 376

```

;
; TYPE: PRT

```

ORGANISM: Homo sapiens

US-10-672-399-4

Query Match 86.5%; Score 1612; DB 16; Length 376;
Best Local Similarity 98.7%; Pred. No. 6e-153;
Matches 311: Conservative 1; Mismatches 3; Indels 0

Qy	1	MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFGLYG	60
Db	1	MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFGLYG	60
Qy	61	AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASABA	120
Db	61	AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASABA	120
Qy	121	LLYPRLRLVLMVDGNRAEOLYVDMFREVFADEDPATYVMDGNTHQWPEAAPAAGAVCA	180
Db	121	LLYPRLRLVLMVDGNRAEOLYVDMFREVFADEDPATYVMDGNTHQWPEAAPAAGAVCA	180
Qy	181	GAYREVEAEDPGRLAVEALVTRRCVCVQARGWGKREVMYTAFAKLGDSVDYVYQVCDSDT	240
Db	181	GAYREVEAEDPGRLAVEALVTRRCVCVQARGWGKREVMYTAFAKLGDSVDYVYQVCDSDT	240
Qy	241	RLDPMALLVRLVLEDDPRVGAAGDVRILNPLDSWYSFLSSLRYWVAFNVERACQSYFH	300
Db	241	RLDPMALLVRLVLEDDPRVGAAGDVRILNPLDSWYSFLSSLRYWVAFNVERACQSYFH	300
Qy	301	CVSCISGPLESCPGP	315
Db	301	CVSCISGSLGTPPGP	315

RESULT 3

US-10-042-523-2
; Sequence 2, Application US/10042523
: Publication No. US20020151026A1

APPLICANT: Briakin Michael J

APPLICANT: BILSKIN, MICHAEL J.
TITLE OF INVENTION: Mammalian Hyaluronan Synthases. Nucleic

1. TITLE OF INVENTION: MAMMALIAN HYALURONAN 5-ACIDS AND USES THEREOF

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/042,523
FILING DATE: 19-Oct-2001

—

APPLICATION NUMBER: US/08/635,552A
FILING DATE: 22-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook David P

```

; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-042-523-2

Query Match      86.1%; Score 1605; DB 13; Length 578;
Best Local Similarity 96.0%; Pred.No. 5.5e-152;
Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY    1 MRQODAPKPTPAARRCSGLARRVLTITAFALLIILGLMTWAYAAGVPLASDRYGLLAFGLY 60
Db    1 MRQODAPKPTPAARRCSGLARRVLTITAFALLIILGLMTWAYAAGVPLASDRYGLLAFGLY 60

QY    61 AFLSAHLVAQSIFAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCLASARA 120
Db    61 AFLSAHLVAQSIFAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCLASARA 120

QY    121 LIYPRLARLVLMVDGNRAEDLYMVDMPREVFADDPATYYWPDGNYHQHPPEPAAAGAVGA 180
Db    121 LIYPRLARLVLMVDGNRAEDLYMVDMPREVFADDPATYYWPDGNYHQHPPEPAAAGAVGA 180

QY    181 GAYREVEARDPGRLAEALVRTRRCVCVQRWGKGKEVMYTAFKALGDSVDYYVQCSDST 240
Db    181 GAYREVEARDPGRLAEALVRTRRCVCVQRWGKGKEVMYTAFKALGDSVDYYVQCSDST 240

QY    241 RLDPMALLELVRVLDDPRVGAVGGDVRIINPLDLSWVSFLSSLRYWVAFNVERACQSYFH 300
Db    241 RLDPMALLELVRVLDDPRVGAVGGDVRIINPLDLSWVSFLSSLRYWVAFNVERACQSYFH 300

QY    301 CVSCISGPLESCGPPEHAMPSFL 325
Db    301 CVSCISGPLE---GLYRNNLLQQOFL 321

```

RESULT 4

US-10-672-399-2
; Sequence 2, Application US/10672399
: Publication No. US20050003368A1

APPLICANT: University of Alberta

APPLICANT: UNIVERSITY OF ALABAMA
TITLE OF INVENTION: Cancer Monitoring and Therapeutics

; TITLE OF INVENTION: CANCER
: FILE REFERENCE: A894635US

FILE REFERENCE: HQ403500
CURRENT APPLICATION NUMBER: US/10/672,399

CURRENT FILING DATE: 2003-09-25

; PRIORITY APPLICATION NUMBER: US 60/472,401

```
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-2

Query Match      86.1%; Score 1605; DB 16; Length 578;
Best Local Similarity 96.0%; Pred. No. 5 5e-152;
Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAAGVPLASDRYGLLAFGLYG 60
Db 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAAGVPLASDRYGLLAFGLYG 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 180
Db 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 180
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
QY 241 RLDPMALLELVRLVDEDPGAVGDDVRLINPLDSWVSLSLRYWVAFNVERACQSYFH 300
Db 241 RLDPMALLELVRLVDEDPGAVGDDVRLINPLDSWVSLSLRYWVAFNVERACQSYFH 300
QY 301 CVSCISGPLESCPGPREHAMMPSFL 325
Db 301 CVSCISGPLESCPGPREHAMMPSFL 325
US-09-902-939-4

Query Match      81.1%; Score 1512; DB 10; Length 583;
Best Local Similarity 90.6%; Pred. No. 1.2e-142;
Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3;

QY 3 QODAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAAGVPLASDRYGLLAFGLYGAF 62
Db 2 QODMPKPEARCCSGLARRALTITFALLILGLMTWYAAAGVPLASDRYGLLAFGLYGAF 61
QY 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 117
Db 62 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 121
QY 118 ARALLYPRARLRVLMVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAG 176
Db 122 ARALLYPRTRUKRVLMVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAG 181

; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-2

Query Match      86.1%; Score 1605; DB 16; Length 578;
Best Local Similarity 96.0%; Pred. No. 5 5e-152;
Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAAGVPLASDRYGLLAFGLYG 60
Db 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAAGVPLASDRYGLLAFGLYG 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 180
Db 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 180
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
QY 241 RLDPMALLELVRLVDEDPGAVGDDVRLINPLDSWVSLSLRYWVAFNVERACQSYFH 300
Db 241 RLDPMALLELVRLVDEDPGAVGDDVRLINPLDSWVSLSLRYWVAFNVERACQSYFH 300
QY 301 CVSCISGPLESCPGPREHAMMPSFL 325
Db 301 CVSCISGPLESCPGPREHAMMPSFL 325
US-10-672-399-8

Query Match      69.6%; Score 1298; DB 16; Length 320;
Best Local Similarity 95.1%; Pred. No. 1.6e-121;
Matches 252; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 180
Db 61 LLYPRARLRVLMVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 120
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db 121 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
QY 241 RLDPMALLELVRLVDEDPGAVGDDVRLINPLDSWVSLSLRYWVAFNVERACQSYFH 300
Db 181 RLDPMALLELVRLVDEDPGAVGDDVRLINPLDSWVSLSLRYWVAFNVERACQSYFH 240
QY 301 CVSCISGPLESCPGPREHAMMPSFL 325
Db 241 CVSCISGPLESCPGPREHAMMPSFL 261

RESULT 7
US-10-042-523-3
; Sequence 3, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
```



```
Db          292 RNSLQQFL 300
RESULT 10
US-10-295-027-370
; Sequence 370, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gieh, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 370
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-370

Query Match          42.4%; Score 790.5; DB 15; Length 553;
Best Local Similarity 53.4%; Pred. No. 3.2e-70;
Matches 164; Conservative 39; Mismatches 85; Indels 19; Gaps 5;

QY          21 RRVLTAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRR 80
Db          10 RVVGTSLFALAVLGGILAAVYTGQFIHTEKHLYSFGLYGAILGHLILIQSLFAFLEHRR 69
QY          81 VAAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRABLRVLMVVDGNRA 139
Db          70 MRRAGQALKPSRRGSVALCIAAYQEDPDYLRKCLRSQAQRISFP--DLKVMVMVVDGNRQ 127
QY          140 EDLYVMDMFREVF-ADEDPATYVWDGNVHQPWEPAAGAVGAGAYREVEADPGRLAYEA 198
Db          128 EDAYMLDIFHEVLGGTEQAGFFVWRNSFHEAGEGETEASLQEGMDR-----VRD 176
QY          199 LVTRRCVCVAQRWGKREVMYTAFAKLGDSVDYVQVCDSDTRLDPMALLELRVLDDEDP 258
Db          177 VVRSTFSCINQKQKGGKREVMYTAFAKLGDSVDYIQVCDSDTVLDPACTIEMLRVLEDDP 236
QY          259 RVGAGGVDVRLNPLDLSWVSFLSSLRVYVAFNVERACQSYFHCVSCISGPLESCEPGPREH 318
Db          237 QVGGVGGDVQILNKYDSWISFLSSVRVYVWMAFNVERACQSYFGCVQCISGPL-----GMVRN 292
QY          319 AMPSFL 325
Db          293 SLQQFL 299

RESULT 12
US-09-902-939-5
; Sequence 5, Application US/09902939
; Publication No. US20030087850A1

Db          237 QVGGVGGDVQILNKYDSWISFLSSVRVYVWMAFNVERACQSYFGCVQCISGPL-----GMVRN 292
QY          319 AMPSFL 325
Db          293 SLQQFL 299

US-10-188-832-137
; Sequence 137, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Nataasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 137
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-137

Query Match          42.4%; Score 790.5; DB 15; Length 553;
Best Local Similarity 53.4%; Pred. No. 3.2e-70;
Matches 164; Conservative 39; Mismatches 85; Indels 19; Gaps 5;

QY          21 RRVLTAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRR 80
Db          10 RVVGTSLFALAVLGGILAAVYTGQFIHTEKHLYSFGLYGAILGHLILIQSLFAFLEHRR 69
QY          81 VAAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRABLRVLMVVDGNRA 139
Db          70 MRRAGQALKPSRRGSVALCIAAYQEDPDYLRKCLRSQAQRISFP--DLKVMVMVVDGNRQ 127
QY          140 EDLYVMDMFREVF-ADEDPATYVWDGNVHQPWEPAAGAVGAGAYREVEADPGRLAYEA 198
Db          128 EDAYMLDIFHEVLGGTEQAGFFVWRNSFHEAGEGETEASLQEGMDR-----VRD 176
QY          199 LVTRRCVCVAQRWGKREVMYTAFAKLGDSVDYVQVCDSDTRLDPMALLELRVLDDEDP 258
Db          177 VVRSTFSCINQKQKGGKREVMYTAFAKLGDSVDYIQVCDSDTVLDPACTIEMLRVLEDDP 236
QY          259 RVGAGGVDVRLNPLDLSWVSFLSSLRVYVAFNVERACQSYFHCVSCISGPLESCEPGPREH 318
Db          237 QVGGVGGDVQILNKYDSWISFLSSVRVYVWMAFNVERACQSYFGCVQCISGPL-----GMVRN 292
QY          319 AMPSFL 325
Db          293 SLQQFL 299
```

```
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 552
; TYPE: PRT
; ORGANISM: mouse
US-09-902-939-5

Query Match      42.0%; Score 783.5; DB 10; Length 552;
Best Local Similarity 51.3%; Pred. No. 1.6e-69;
Matches 157; Conservative 52; Mismatches 76; Indels 21; Gaps 7;

QY 21 RRVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHR 79
DB 11 RIIGTTLFGVSLLLGITAAYIVGYQFIQTDNY-YFSFGLYGAFGLASHLIIQSLFAFLHR 69
QY 80 RVAARAGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRA 139
DB 70 KMKKSLETPI--KLKNTVALCIAAQEDPDYLRKLCQSVKRLTYP--GIKVMWIDGNSD 125
QY 140 EDLYMVMDFREVFADDEDPATYVWDGNVGHQWPEPAAAGAVGAGAYREVEAEDPGRLAVAL 199
DB 126 DDLYMMDIFSEVIGRDKSATYIWKNNFHEK-----GPGTEESHKESQH--VTQL 174
QY 200 VTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDTFLDPMALLELRVLDEPR 259
DB 175 VLSNKSICIMQKGGKREVMYTAFAKALGDSVDYVQVCDSDTFLDPMALLELRVLDEPR 234
QY 260 VGAVGGDVRIINPLDWSVFLSSRLRYWAFNVERACQSYFHCVSCISGPSCPGPREHA 319
DB 235 VGGVGGDVQILNKYDSWISFLSSRYWAFNVERACQSYFHCVSCISGPSCPGPREHA 319
QY 320 MMPSFL 325
DB 291 LLHEFV 296

Query Match      42.0%; Score 783.5; DB 14; Length 552;
Best Local Similarity 51.0%; Pred. No. 1.6e-69;
Matches 157; Conservative 53; Mismatches 77; Indels 21; Gaps 7;

QY 19 LARRVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFGLYGAFLSAHLVAQSLFAYLE 77
DB 9 ILRIIGTTLFGVSLLLGITAAYIVGYQFIQTDNY-YFSFGLYGAFGLASHLIIQSLFAFL 67

; GENERAL INFORMATION:
; APPLICANT: Yuichi Wada
; APPLICANT: Hideshige Moriya
; TITLE OF INVENTION: GENES FOR TRANSFECTION INTO BONY TISSUES
; FILE REFERENCE: OPI418
; CURRENT APPLICATION NUMBER: US/10/262,526
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: JP 2001-367091
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-262-526-4

Query Match      42.0%; Score 783.5; DB 14; Length 552;
Best Local Similarity 51.3%; Pred. No. 1.6e-69;
Matches 157; Conservative 52; Mismatches 76; Indels 21; Gaps 7;

QY 21 RRVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHR 79
DB 11 RIIGTTLFGVSLLLGITAAYIVGYQFIQTDNY-YFSFGLYGAFGLASHLIIQSLFAFLHR 69
QY 80 RVAARAGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRA 139
DB 70 KMKKSLETPI--KLKNTVALCIAAQEDPDYLRKLCQSVKRLTYP--GIKVMWIDGNSD 125
QY 140 EDLYMVMDFREVFADDEDPATYVWDGNVGHQWPEPAAAGAVGAGAYREVEAEDPGRLAVAL 199
DB 126 DDLYMMDIFSEVIGRDKSATYIWKNNFHEK-----GPGTEESHKESQH--VTQL 174
QY 200 VTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDTFLDPMALLELRVLDEPR 259
DB 175 VLSNKSICIMQKGGKREVMYTAFAKALGDSVDYVQVCDSDTFLDPMALLELRVLDEPR 234
QY 260 VGAVGGDVRIINPLDWSVFLSSRLRYWAFNVERACQSYFHCVSCISGPSCPGPREHA 319
DB 235 VGGVGGDVQILNKYDSWISFLSSRYWAFNVERACQSYFHCVSCISGPSCPGPREHA 319
QY 320 MMPSFL 325
DB 291 LLHEFV 296

RESULT 13
US-10-262-526-2
; Sequence 2, Application US/10262526
; Publication No. US20030108531A1
; GENERAL INFORMATION:
; APPLICANT: Hideshige Moriya
; APPLICANT: Yuichi Wada
; TITLE OF INVENTION: GENES FOR TRANSFECTION INTO BONY TISSUES
; FILE REFERENCE: OPI418
; CURRENT APPLICATION NUMBER: US/10/262,526
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: JP 2001-367091
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-526-2

Query Match      42.0%; Score 783.5; DB 14; Length 552;
Best Local Similarity 51.0%; Pred. No. 1.6e-69;
Matches 157; Conservative 53; Mismatches 77; Indels 21; Gaps 7;

QY 19 LARRVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFGLYGAFLSAHLVAQSLFAYLE 77
DB 9 ILRIIGTTLFGVSLLLGITAAYIVGYQFIQTDNY-YFSFGLYGAFGLASHLIIQSLFAFL 67

; GENERAL INFORMATION:
; Sequence 4, Application US/10262526
; Publication No. US20030108531A1
; GENERAL INFORMATION:
; APPLICANT: Hideshige Moriya
; APPLICANT: Yuichi Wada
; TITLE OF INVENTION: GENES FOR TRANSFECTION INTO BONY TISSUES
; FILE REFERENCE: OPI418
; CURRENT APPLICATION NUMBER: US/10/262,526
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: JP 2001-367091
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-262-526-4

Query Match      42.0%; Score 783.5; DB 14; Length 552;
Best Local Similarity 51.3%; Pred. No. 1.6e-69;
Matches 157; Conservative 52; Mismatches 76; Indels 21; Gaps 7;

QY 78 HRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGN 137
DB 68 HRMKKSLETPI--KLKNTVALCIAAQEDPDYLRKLCQSVKRLTYP--GIKVMWIDGN 123
QY 138 RAEDLYMVMDFREVFADDEDPATYVWDGNVGHQWPEPAAAGAVGAGAYREVEAEDPGRLAVAL 197
DB 124 SEDDLNMDIFSEVIGRDKSATYIWKNNFHEK-----GPGTEESHKESQH--VT 172
QY 198 ALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDTFLDPMALLELRVLDE 257
DB 173 QLVLSNKSICIMQKGGKREVMYTAFAKALGDSVDYVQVCDSDTFLDPMALLELRVLDE 232
QY 258 PRVGAAGDVRIINPLDWSVFLSSRLRYWAFNVERACQSYFHCVSCISGPSCPGPRE 317
DB 233 PMVGGVGGDVQILNKYDSWISFLSSRYWAFNVERACQSYFHCVSCISGPSCPGPRE 288
QY 318 HAMPSFL 325
DB 289 NSLLHEFV 296

RESULT 14
US-10-262-526-4
; Sequence 4, Application US/10262526
; Publication No. US20030108531A1
; GENERAL INFORMATION:
; APPLICANT: Hideshige Moriya
; APPLICANT: Yuichi Wada
; TITLE OF INVENTION: GENES FOR TRANSFECTION INTO BONY TISSUES
; FILE REFERENCE: OPI418
; CURRENT APPLICATION NUMBER: US/10/262,526
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: JP 2001-367091
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-262-526-4

Query Match      42.0%; Score 783.5; DB 14; Length 552;
Best Local Similarity 51.3%; Pred. No. 1.6e-69;
Matches 157; Conservative 52; Mismatches 76; Indels 21; Gaps 7;

QY 21 RRVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHR 79
DB 11 RIIGTTLFGVSLLLGITAAYIVGYQFIQTDNY-YFSFGLYGAFGLASHLIIQSLFAFLHR 69
QY 80 RVAARAGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRA 139
DB 70 KMKKSLETPI--KLKNTVALCIAAQEDPDYLRKLCQSVKRLTYP--GIKVMWIDGNSD 125
QY 140 EDLYMVMDFREVFADDEDPATYVWDGNVGHQWPEPAAAGAVGAGAYREVEAEDPGRLAVAL 199
DB 126 DDLYMMDIFSEVIGRDKSATYIWKNNFHEK-----GPGTEESHKESQH--VTQL 174
QY 200 VTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDTFLDPMALLELRVLDEPR 259
DB 175 VLSNKSICIMQKGGKREVMYTAFAKALGDSVDYVQVCDSDTFLDPMALLELRVLDEPR 234
QY 260 VGAVGGDVRIINPLDWSVFLSSRLRYWAFNVERACQSYFHCVSCISGPSCPGPREHA 319
DB 235 VGGVGGDVQILNKYDSWISFLSSRYWAFNVERACQSYFHCVSCISGPSCPGPREHA 319
QY 320 MMPSFL 325
DB 291 LLHEFV 296

RESULT 15
US-09-879-959-7
```

```
; Sequence 7, Application US/09879959
; Patent No. US20020160489A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kehama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS EQUISIMILIS HYALURONAN SYNTHASE GENE AND EXPRESSION
; FILE REFERENCE: 3554.049
; CURRENT APPLICATION NUMBER: US/09/879,959
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Chlorella virus PBCV-1
US-09-879-959-7

Query Match      16.6%; Score 310; DB 9; Length 567;
Best Local Similarity 29.4%; Pred. No. 5.8e-22;
Matches 89; Conservative 57; Mismatches 107; Indels 50; Gaps 10;

QY 19 LARRVLTITAFALLIIG-----LMTWAVAAGVPLASDRYGLLAFGLYGAFLSAHLVAOSL 72
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 15 ITSNIIVGGASLIILAPAITGYVLHWNIALSTINGVSAYGIFVGF-----LAQVL 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 73 FAYLEHRR-----VAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARL 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 67 FSELNRKRLRWISLRPKGNDV----RLAVIAGYREDPYWFKCLESVRDSYGNV-A 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 129 RVLMTVDGNRAEDLYWDMFREVFADEDPATYWDGNTHQWPAPAAAGAGAYREVEA 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 122 RLICVIDGDEDDMRMAAVYKAIYND-----NIKKP-----EFVLCE 159
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 189 EDPGRLAVEALVTRRCVCVQRCQWKGKREVMYTFK--ALGDSVDYVQVCDSDTRLDPMA 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 160 DKEGERIDS--DFSRDICVLQPHRGKRECLTGFQLAKMDPSVNAVLIIDSDTVLEKDA 217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 247 LLELYRVLDEDPVAGAGDVRIILNPLDSWSVSLSLRYWAFNVERACQSYFHCVCSCIS 306
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 218 ILEVYVPLACDPEIQAVAGECKIWN-TDTLLSLLVAMRYSAFCVSAQSFRTVQCVG 276
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 307 GPL 309
DB 277 GPL 279
```

Search completed: March 11, 2005, 14:44:38
Job time : 81.7955 secs

THIS PAGE BLANK (SP10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 02:13:47 ; Search time 5030.06 Seconds
(without alignments)
10259.279 Million cell updates/sec

Title: US-10-672-399-7
Perfect score: 1065
Sequence: 1 atgagacagcagcgcgc.....tgggttatgtaccasgtaa 1065

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ets.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1064	99.9	2088	9 HSU59269	US9269 Human hyalu
2	1064	99.9	2116	6 AR220003	AR220003 Sequence
3	1051.8	98.8	2119	9 BC035837	BC035837 Homo sapi
4	1051.6	98.7	2108	9 HUM845	D94424 Homo sapien
5	1051.6	98.7	2117	6 AR137038	AR137038 Sequence
6	986.2	92.6	2109	9 AY463695	AY463695 Papio anu
7	791.4	74.3	2095	10 AB097568	AB097568 Rattus no
8	786.6	73.9	2102	6 E13681	E13681 DNA encodin
9	786.6	73.9	2102	6 E28454	E28454 Hyaluronate
10	786.6	73.9	2102	6 E30971	E30971 Hyaluronon
11	786.6	73.9	2102	6 E34326	E34326 DNA for gen
12	786.6	73.9	2102	10 MUSHAS	D82964 Mus muscucu
13	751.4	70.6	1581	6 CQ731079	CQ731079 Sequence
14	697.2	65.5	229155	9 AC018755	AC018755 Homo sapi
15	669	62.8	193986	2 AC130783	AC130783 Pan trogl
16	635.4	59.7	185623	2 AC137058	AC137058 Papio anu
17	492	46.2	293184	2 AC079487	AC079487 Mus muscu
18	490.8	46.1	72955	2 AC108651	Continuation (4 of
19	490.8	46.1	257325	2 AC116203	AC116203 Rattus no

20	401.2	37.7	1375	5 AY437407	AY437407 Danio rer
C 21	369.2	34.7	611	11 BV099313	BV099313 RPAMSEQ0
C 22	368.2	34.6	611	11 BV160501	BV160501 RPAMSEQ0
23	365	34.4	1659	4 AB055979	AB055979 Oryctolag
24	353	33.1	1682	6 CQ723590	CQ723590 Sequence
25	353	33.1	1994	9 AF234839	AF234839 Homo sapi
26	353	33.1	4049	9 AF232772	AF232772 Homo sapi
27	353	33.1	4220	6 CQ776471	CQ776471 Sequence
28	347.2	32.6	3995	4 AB159675	AB159675 Sus scrof
29	325.4	30.6	2093	5 BC047963	BC047963 Xenopus l
30	323.8	30.4	2950	5 XELDG42A	M22249 X.laavis DG
31	320.2	30.1	1465	10 AB097569	AB097569 Rattus no
32	311.6	29.3	1665	6 E30973	E30973 Hyaluronon
33	311.6	29.3	1665	6 AR373269	AR373269 Sequence
34	311.6	29.3	5919	6 CQ777380	CQ777380 Sequence
35	311.6	29.3	5919	10 MMU86408	U86408 Mus muscucu
36	299.8	28.2	2234	5 AF015780	AF015780 Xenopus l
37	296.6	27.8	2245	5 BC077983	BC077983 Xenopus l
38	274.2	25.7	550	11 BV161377	BV161377 RPAMSEQ0
39	252	23.7	2280	5 AY302252	AY302252 Xenopus l
40	249	23.4	1180	10 AF543196	AF543196 Rattus no
C 41	246.8	23.2	1653	6 BD133837	BD133837 Malignant
42	246.8	23.2	2938	6 E30972	E30972 Hyaluronon
43	246.8	23.2	2947	6 AR285323	AR285323 Sequence
44	246.8	23.2	2947	6 AR373260	AR373260 Sequence
45	246.8	23.2	2948	6 AX305990	AX305990 Sequence

ALIGNMENTS

RESULT 1
HSU59269
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
PUBMED
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

HSU59269 2088 bp mRNA linear PRI 24-SEP-1996
Human hyaluronan synthase mRNA, complete cds.
U59269.1 GI:1556464
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2088)
Shyjan,A.M., Heldin,P., Butcher,E.C., Yoshino,T. and Briskin,M.J.
Functional cloning of the cDNA for a human hyaluronan synthase
J. Biol. Chem. 271 (38), 23395-23399 (1996)
96394438
8798544
2 (bases 1 to 2088)
Briskin,M.J. and Shyjan,A.M.
Direct Submission
Submitted (24-MAY-1996) LeukoSite Inc., 215 First Street,
Cambridge, MA 02142, USA
Location/Qualifiers
1..2088
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
36..1772
/codon_start=1
/product="hyaluronan synthase"
/protein_id="AAC50706.1"
/db_xref="GI:1556465"
/translation="MRQQDAPKPTPAARCSGLARRVLTIAFALLIILMTWAYAGV
PLASDRIGLLAFGLYGAFLSAHLVAQSLFAYLEHRRVAAARGLDRAATARSVALTIS
AYQEDPAYLRQLCLASARALLYPRALRLVMDGNRAEDLYMDMREVPADDDPATY
VMDGNTHQWPEAAGAGVAGAYREVEADPGRLAVEALVTRRCVCVCAQORWGKREV
MYTAFKALGSDYDVQVDSIDRLDPMALLELVRLDDEPRVAGVDVRLILPLDSW
VSPFLSLRYWVAFNVERACQSYCHVCISGPIGLYRNLLQOFLAANYNQKFLGTHC
TFGDDRLHTRMLSMGYATKYTSRSCYSETSPSLRWLSQQTRWKSQYREWLNNAL
WHRHHAMWTYEAIVSGLFFVFAATVRLFYAGRWALLVLLCVQGVAKAAAFPA
WLRGCUWVLLSLYAPLYMCGLLPAKFLALVTMNGSGWTSGRKLAANYVPLPLAL
WALLLLGLVRSVAHEARADWSGFSRAAEAYHLAAGAGAYVGYWVAAMLTLVYWGVRRL

Db	996	CTTGAGGCGCTGTGATCAACAGAGTTCCTGGGTACCCACTGTACTCTTTTGGGGATGACCGG	1055
QY	1021	CACCTCACCAACCGCATGCTCAGCATGGGTTATGCTACCAAGTA	1064
Db	1056	CACCTCACCAACCGCATGCTCAGCATGGGTTATGCTACCAAGTA	1099
RESULT 2			
AR220003			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
/organism= "unknown"			
/mol_type= "mRNA "			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1064; Conservative			
QY	1	ATGAGACAGCAGGACGCGCCCAAGCCCACTCCTCGAGCCCGCCGCTGCTCGGCTTGCC	60
Db	36	ATGAGACAGCAGGACGCGCCCAAGCCCACTCCTCGAGCCCGCCGCTGCTCGGCTTGCC	95
QY	61	CGAGAGGTGTGACCATCGCTTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC	120
Db	96	CGAGAGGTGTGACCATCGCTTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC	155
QY	121	GCGCGCGGGGTGCGCTGGCGCTTCGATCGCTACGGCTCTCTGGGCTTCGGGCTTACGGG	180
Db	156	GCGCGCGGGGTGCGCTGGCGCTTCGATCGCTACGGCTCTCTGGGCTTCGGGCTTACGGG	215
QY	181	GCGCTTCCTTTACGGCAGCTGTGGCGAGGCTCTTCGGGTACCTGGAGCAGCGGCGG	240
Db	216	GCGCTTCCTTTACGGCAGCTGTGGCGAGGCTCTTCGGGTACCTGGAGCAGCGGCGG	275
QY	241	GTGGCGGCGCGCGCGCGCTGGATCGATCGATCGATCGATCGATCGATCGATCG	300
Db	276	GTGGCGGCGCGCGCGCGCTGGATCGATCGATCGATCGATCGATCGATCGATCG	335
QY	301	ATCTCCGCTTCAAGCGCTTCGAGATTCGGTGGATCTGTCAGGCTCTGTCAGGACACA	360
Db	336	ATCTCCGCTTCAAGCGCTTCGAGATTCGGTGGATCTGTCAGGCTCTGTCAGGACACA	395
QY	361	GTGGCGGCGCGCGCGCGCTGGATCGATCGATCGATCGATCGATCGATCGATCG	420
Db	396	GTGGCGGCGCGCGCGCGCTGGATCGATCGATCGATCGATCGATCGATCGATCG	455
QY	421	GACCTCTACATGGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCACGTAC	480
Db	456	GACCTCTACATGGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCACGTAC	515
QY	481	GTGTGGGACGGCAACTACACAGCCCTCGGGAACCCCGCGCGCGCGGCTGGGCGCC	540
Db	516	GTGTGGGACGGCAACTACACAGCCCTCGGGAACCCCGCGCGCGCGGCTGGGCGCC	575
QY	541	GGAGCTTATCGGAGGTGGAGGCGGAGGATCTCTGGGCGGCTGGCAGTGGAGGCGCTGGTG	600
Db	576	GGAGCTTATCGGAGGTGGAGGCGGAGGATCTCTGGGCGGCTGGCAGTGGAGGCGCTGGTG	635
QY	601	AGGACTTCGAGGTGCGTGTGGCGAGCGCTGGGCGGCAAGCGCGAGGTTCATGTAC	660
Db	636	AGGACTTCGAGGTGCGTGTGGCGAGCGCTGGGCGGCAAGCGCGAGGTTCATGTAC	695
QY	661	ACAGCCTTCAAGCGCTTCGAGATTCGGTGGATCTGTCAGGCTCTGTCAGGACACA	720
Db	696	ACAGCCTTCAAGCGCTTCGAGATTCGGTGGATCTGTCAGGCTCTGTCAGGACACA	755
QY	721	AGGTTGGAACCCCATGGCACTGCTGGAGCTCTGGGCGGCTGGAGGAGGACCCCGGGTA	780
Db	756	AGGTTGGAACCCCATGGCACTGCTGGAGCTCTGGGCGGCTGGAGGAGGACCCCGGGTA	815
QY	781	GGGGCTGTGGTGGGACGTGGGATCCTTAACTCTGAGCTCTCTGGGTCAGCTTCCTTA	840
Db	816	GGGGCTGTGGTGGGACGTGGGATCCTTAACTCTGAGCTCTCTGGGTCAGCTTCCTTA	875
QY	841	AGCAGCTTCGATCTAGGCTAGCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCAC	900
Db	876	AGCAGCTTCGATCTAGGCTAGCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCAC	935
QY	901	TGTGTATCTGATCAGCGGCTCTTAGGCTTATATAGGAATAACCTCTTGGAGAGTTT	960
Db	936	TGTGTATCTGATCAGCGGCTCTTAGGCTTATATAGGAATAACCTCTTGGAGAGTTT	995
QY	961	CTTGAGGCGCTGTGATCAACAGAGTTCCTGGGTACCCACTGTACTCTTTTGGGGATGACCGG	1020

Db 636 AGGACTCGCAGTGTGCTGTGGCGCAGCGCTGGGGCGGCAAGCGGAGTCAATGTAC 695
Qy 661 ACAGCGCTTCAAGCGCGCTCGGAGATTCCGTGGACTACGTGCAGGTCTGTGACTCGGACACA 720
Db 696 ACAGCGCTTCAAGCGCGCTCGGAGATTCCGTGGACTACGTGCAGGTCTGTGACTCGGACACA 755
Qy 721 AGGTTGGACCCCATGCTGCTGAGCTGTGCGGTTACTGGACGAGACCCCGGGTGA 780
Db 756 AGGTTGGACCCCATGCTGCTGAGCTGTGCGGTTACTGGACGAGACCCCGGGTGA 815
Qy 781 GGGGCTGTGGTGGGAGCTGGGATCCCTTAACCTCTGGAATCCTGGGTGAGTTCCTA 840
Db 816 GGGGCTGTGGTGGGAGCTGGGATCCCTTAACCTCTGGAATCCTGGGTGAGTTCCTA 875
Qy 841 AGCAGCTCGGATCTGCTGCTGAGCTTCAATGTGGAGCGGGCTTGTGACAGTCTACTTCCAC 900
Db 876 AGCAGCTCGGATCTGCTGCTGAGCTTCAATGTGGAGCGGGCTTGTGACAGTCTACTTCCAC 935
Qy 901 TGTGTATCTGCTATCAGCGGCTCTCTAGGCTTATATAGGAATACTCTTGCAGGAGTTT 960
Db 936 TGTGTATCTGCTATCAGCGGCTCTCTAGGCTTATATAGGAATACTCTTGCAGGAGTTT 995
Qy 961 CTGTAGGCTGTGACACAGAGTTCTGGGTACCCAGTGTACTTTCGGGATGACCGG 1020
Db 996 CTGTAGGCTGTGACACAGAGTTCTGGGTACCCAGTGTACTTTCGGGATGACCGG 1055
Qy 1021 CACCTCACCAACCGCATGCTGAGCATGGGTATGCTTACCAAGTA 1064
Db 1056 CACCTCACCAACCGCATGCTGAGCATGGGTATGCTTACCAAGTA 1099

BC035837 2119 bp mRNA linear PRI 30-JUN-2004
Homo sapiens hyaluronan synthase 1, mRNA (cDNA clone MGC:46218
IMAGE:589083), complete cds.

ACCESSION BC035837
VERSION BC035837.1 GI:23243101
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2119)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Caeavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalius, D.E.,
Schnurch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2119)
Straussberg, R.
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Ahter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, J., Legaspi, R.,
Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, B.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAK Plate: 79 Row: f Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504338.

FEATURES

source

1..2119
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:46218 IMAGE:589083"
/tissue_type="Ovary, pooled from 3 adults"
/clone_lib="NIH MGC_125"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1..2119
/gene="HAS1"
/note="synonym: HAS"
/db_xref="LocusID:3036"
/db_xref="MIM:601463"
42..1775
/gene="HAS1"
/codon_start=1
/product="hyaluronan synthase 1"
/protein_id="AAH35837.1"
/db_xref="GI:23243102"
/db_xref="LocusID:3036"
/db_xref="MIM:601463"
/translation="MRQAPKPTPAARCSGLARRVLTIFALLILGLMTWYAAAGVP
LASDRYGLLAFLGALFSLAHVAQSLPAYLEHRRVAAARGLDAAATVALTISA
YQEDPAYLRQCLAGARALLYPRARLVLMVVDGNRAEDLYMVDNFREVFADBDPATYV
WDGNYPWPEPAAGAGAGAVREVEDPGLRAVEALVTRRCVCVAORWGGKREVM
YTFKALGDSNIVQVCDSTRLDPMALLELVRLVDEPRVGVGVDVRLNPLDSHW
SFLSLRYWAFNVERACSYFHCVCISGIGLYRNLLQQLFLEAWYNQKPLGTHCT
FGDDRHLNRLMSGYATKYSRSCYSETSSFLWLSQQTRFWSKFRBWLYNALW
WHRHAWMTYEAIVSGIFPFVAAATVRLFYAGRPWALLVLLCVQGVALLAAAPAAW
LRGLCRMVLLSLYAPLMCGLLPAKFLALVTMNSGCTSGRRKLAANVPLPLALW
ALLGLGIVRSVAHEARADWSGPRABAYHLAAGAYGVGVWVWMLTYWGVVRRLC
RRRTGGYRVQV"

gene

CDS

ORIGIN

Query Match 98.8%; Score 1051.8; DB 9; Length 2119;
Best Local Similarity 99.8%; Pred. No. 1.2e-134;
Matches 1053; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 10 CAGGACGCGCCCAAGCCCACTCTTCACACCCCGCGCTGCTCGGCTGCGCCGCGGGTG 69
Db 48 CAGGACGCGCCCAAGCCCACTCTTCACACCCCGCGCTGCTCGGCTGCGCCGCGGGTG 107
Qy 70 CTGACCATCGCTTCGCGCTGCTCATCTCGGCGCTCATGACCTGGCGCTACGCCCGGG 129
Db 108 CTGACCATCGCTTCGCGCTGCTCATCTCGGCGCTCATGACCTGGCGCTACGCCCGGG 167

QY	130	GTGCGCTGGCTCGATCGCTACCGGCTCTCGGCTTCTCGGCTCTACGGGGCTTCCTT	189
Db	168	GTGCGCTGGCTCGATCGCTACCGGCTCTCGGCTTCTCGGCTCTACGGGGCTTCCTT	227
QY	190	TACAGCGCACTGCTGGCGCAGAGCTCTTTCGCTACCTGGAGCACCGGCGGTGCGGCG	249
Db	228	TCAGCGCACTGCTGGCGCAGAGCTCTTTCGCTACCTGGAGCACCGGCGGTGCGGCG	287
QY	250	GGGCGCGGGGCGCTGTGATGTCAGCACCGCGCGAGTGTGGGCTGACCAATCTCGCC	309
Db	288	GGGCGCGGGGCGCTGTGATGTCAGCACCGCGCGAGTGTGGGCTGACCAATCTCGCC	347
QY	310	TACAGAGGACCGCGCTACCTGCGCAGTGCCTGGGCTCGCGCGCGCTGCTGTAC	369
Db	348	TACAGAGGACCGCGCTACCTGCGCAGTGCCTGGGCTCGCGCGCGCTGCTGTAC	407
QY	370	CGCGCGCGGCTGCGGCTCTCATGTGTGTGATGTCACCGCGCGAGGACCTCTAC	429
Db	408	CGCGCGCGGCTGCGGCTCTCATGTGTGTGATGTCACCGCGCGAGGACCTCTAC	467
QY	430	ATGCTCGACATGTTCCGAGAGTCTTCGCTGACAGGACCCCGCCACATGATGTGGAC	489
Db	468	ATGCTCGACATGTTCCGAGAGTCTTCGCTGACAGGACCCCGCCACATGATGTGGAC	527
QY	490	GGCACTACCAACCGGCTGGAACCCCGCGCGGCGGCGGCTGGGCGCGGAGCTAT	549
Db	528	GGCACTACCAACCGGCTGGAACCCCGCGCGGCGGCGGCTGGGCGCGGAGCTAT	587
QY	550	CGGAGGTGGAGCGGAGATCCTGGGCGGCTGGCAGTGGAGCGCTGCTGAGGACTCGC	609
Db	588	CGGAGGTGGAGCGGAGATCCTGGGCGGCTGGCAGTGGAGCGCTGCTGAGGACTCGT	647
QY	610	AGTGCCTGTGCTGGCGAGCGCTGGGCGGCGCAAGCGCGAGTCAATGACAGCCTTC	669
Db	648	AGTGCCTGTGCTGGCGAGCGCTGGGCGGCGCAAGCGCGAGTCAATGACAGCCTTC	707
QY	670	AAGCGCTCGAGATTCGCTGAGCTACGTCAGAGTCTGCTGCGGACACAGGTTGGAC	729
Db	708	AAGCGCTCGAGATTCGCTGAGCTACGTCAGAGTCTGCTGCGGACACAGGTTGGAC	767
QY	730	CCATGCGACTGCTGGAGCTCTGGGCGGCTACGTCAGAGGACCCCGCGGTAGGGCTGT	789
Db	768	CCATGCGACTGCTGGAGCTCTGGGCGGCTACGTCAGAGGACCCCGCGGTAGGGCTGT	827
QY	790	GGTGGGACGTCGAGCTTAACCCCTCTGAGCTCTGGGCTGAGCTTCTTAAGCAGCCTG	849
Db	828	GGTGGGACGTCGAGCTTAACCCCTCTGAGCTCTGGGCTGAGCTTCTTAAGCAGCCTG	887
QY	850	CGATACGCTGAGCTTCAATGTGGAGCGGCTTGTGAGGCTACTTCCACTGTATCC	909
Db	888	CGATACGCTGAGCTTCAATGTGGAGCGGCTTGTGAGGCTACTTCCACTGTATCC	947
QY	910	TGCATCAGCGTCTCTAGGCTATATAGGAATAACCTCTTGAGCAGTCTTCTTGAGGCC	969
Db	948	TGCATCAGCGTCTCTAGGCTATATAGGAATAACCTCTTGAGCAGTCTTCTTGAGGCC	1007
QY	970	TGGTACACAGAGTTCCTGGGTACCACTGTACTTTTGGGATGACCGGCACTCACC	1029
Db	1008	TGGTACACAGAGTTCCTGGGTACCACTGTACTTTTGGGATGACCGGCACTCACC	1067
QY	1030	AACCGCATGCTCAGCATGGTATGCTACCAAGTA	1064
Db	1068	AACCGCATGCTCAGCATGGTATGCTACCAAGTA	1102

RESULT 4
HUMHAS HUMHAS 2108 bp mRNA linear PRI 06-FEB-1999
LOCUS Homo sapiens mRNA for hyaluronan synthase, complete cds.
DEFINITION D84424
ACCESSION D84424.1 GI:1401033
VERSION hHAS; hyaluronan synthase.
KEYWORDS Homo sapiens (human)
SOURCE

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 2108)
TITLE	Itano.N. and Kinata.K.
JOURNAL	Molecular cloning of human hyaluronan synthase
MEDLINE	Biochem. Biophys. Res. Commun. 222 (3), 816-820 (1996)
PUBMED	96244584
REFERENCE	2 (bases 1 to 2108)
AUTHORS	Itano.N.
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 2108)
AUTHORS	Itano.N.
TITLE	Direct Submission
JOURNAL	Submitted (16-APR-1996) Naoki Itano, Aichi Medical University, Institute for Molecular Science of Medicine; Nagakute, Aichi-gun, Aichi 480-11, Japan (Tel:052-264-4811(ex.2087), Fax:0561-63-3532)
FEATURES	Location/Qualifiers
source	1..2108
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/tissue_type="brain"
	/clone_lib="lg11"
	/dev_stage="fetal"
gene	1..2108
	/gene="hHAS"
CDS	149..1780
	/gene="hHAS"
	/codon_start=1
	/product="hyaluronan synthase"
	/protein_id="BAAL2351.1"
	/translation="MTWAYAAGVPLASDRYGLLAFGLYAFSLVAHQAQSLFAYLEHR
	RVAAAGSLDAATARSVALTISAYQDPALVROCLASARALLYPRARVRVLMVDGN
	RAEDLVMDPREVEFADEDPATVYMDGNVHPWEPAAAGVAGAGAYREVEADEPRLA
	VEALVTRRCVCVAQRWGGKREVMYAFKALGDSVDVQVCDSTRDLPMALLLRV
	LDEPRVAGGVDRILNPLDSWVSFLUSLRVAFNVRACQSFHCVCISGLPLG
	YRNLLQOFLAEMYNQFLGTHCTFGDDRHLTNRMLSMGYATKTSRSCYSETPSSP
	LRWLSQOTRWKSKYFRELWNALMHRHAWMTVEAVSGLFPFFVAATVLRLLFVAGQ
	PWALLWLLCVQVALAKAAPAWLRGLRMVLLSLVAPLYMCLLPKPTLALVTMNO
	SWGTSGRRLAANYVPLPLALWALLGLGLVPSVAHEARADWSGFSRAEAHYLAA
	GAGAYGVWAMLTLYWGVRLRRTTGGTVQV"
ORIGIN	
Query Match	98.7%; Score 1051.6; DB 9; Length 2108;
Best Local Similarity	99.6%; Pred. No. 1.3e-134;
Matches 1054; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
QY	7 CAGCAGACGCGCCCAAGCCACTCTCTGCAGCCCGCGCTGCTCGGCTGGCCGAGG 66
Db	50 CAGCAGACGCGCCCAAGCCACTCTCTGCAGCCCGCGCTGCTCGGCTGGCCGAGG 109
QY	67 GTGCTGACCATCGCTTTCGCCCTGCTCATCTGGGCTCATGACTGGGCTACGCCCC 126
Db	110 GTGCTGACCATCGCTTTCGCCCTGCTCATCTGGGCTCATGACTGGGCTACGCCCC 169
QY	127 GGGGTGCGCTGGCTCCGATCGCTAGCGCTCTGGGCTTCGGGCTTACGGGCTTTC 186
Db	170 GGGGTGCGCTGGCTCCGATCGCTAGCGCTCTGGGCTTCGGGCTTACGGGCTTTC 229
QY	187 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTTCGGTACTGCTGGAGACCGCGCGGTGGC 246
Db	230 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTTCGGTACTGCTGGAGACCGCGCGGTGGC 289
QY	247 GCGGCGCGGGGGCGCTGGATGACGACCGCGCGAGTGGGCTGACCATCTCC 306
Db	290 GCGGCGCGGGGGCGCTGGATGACGACCGCGCGAGTGGGCTGACCATCTCC 349
QY	307 GCCTACAGGAGACCCCGGCTACTCTGGCAGTGCCTGGCGCTCCGCCCGCGCTGCTG 366
Db	350 GCCTACAGGAGACCCCGGCTACTCTGGCAGTGCCTGGCGCTCCGCCCGCGCTGCTG 409

Matches 1054; Conservative		0; Mismatches	4; Indels	0; Gaps	0
Qy	7	CAGCAGACGCGCCAAAGCCCACTCTCGACGCCGCGCTGCTCCGGCCTGGCCCGGAGG	66		
Db	50	CAGCAGACGCGCCAAAGCCCACTCTCTGCAGCCGCGCTGCTCGGCCCTGGCCCGGAGG	109		
Qy	67	GTGCTGACCATCGCCTTGGCCCTGCTCATCTTGGGCCCTCATGACCTGGGCCCTACGCCGC	126		
Db	110	GTGCTGACCATCGCCTTGGCCCTGCTCATCTTGGCCCTCATGACCTGGGCCCTACGCCGC	169		
Qy	127	GGGCTGCGCTGGCCTCCGATCGCTACGGCTCTCTGGCCTTTCGGCCTCTACGGGCGCTTC	186		
Db	170	GGGCTGCGCTGGCCTCCGATCGCTACGGCTCTCTGGCCTTTCGGCCTCTACGGGCGCTTC	229		
Qy	187	CTTTTCAGCGCACTGTGTGGCGCAGAGCCTCTTTCGCTACCTTGGAGCACCGCGGTGGCG	246		
Db	230	CTTTTCAGCGCACCTGTGTGGCGCAGAGCCTCTTTCGCTACCTTGGAGCACCGCGGTGGCG	289		
Qy	247	GCGCGCGCGCGGGCCGCTGGATGACAGCCACACCGCGCGCAGTGTGGCGCTGACCATCTCC	306		
Db	290	GCGCGCGCGCGGGCCGCTGGATGACAGCCACACCGCGCGCAGTGTGGCGCTGACCATCTCC	349		
Qy	307	GCCTTACCAGAGAGACCCCGGTACTCTCGCGCAGTGCCTGCGCTCCGCCCGGCCCTCGTG	366		
Db	350	GCCTTACCAGAGAGACCCCGGTACTCTCGCGCAGTGCCTGCGCTCCGCCCGGCCCTCGTG	409		
Qy	367	TACCCGCGCGCGGCTGCGCGTCTCATGTGTGTGATGCCAACCGCGCGCGAGCACTC	426		
Db	410	TACCCGCGCGCGGCTGCGCGTCTCATGTGTGTGATGCCAACCGCGCGCGAGCACTC	469		
Qy	427	TACATGTTGACATGTTTCCGCGAGGTTCTCGCTGAACGAGCACCCGCCACGTTAGTGTGG	486		
Db	470	TACATGTTGACATGTTTCCGCGAGGTTCTCGCTGAACGAGCACCCGCCACGTTAGTGTGG	529		
Qy	487	GAGGCGAACTACCAACAGCCCTGGGACCCGCGCGCGCGCGCGGTGGCGCGCGGAGCC	546		
Db	530	GAGGCGAACTACCAACAGCCCTGGGACCCGCGCGCGCGCGCGGTGGCGCGCGGAGCC	589		
Qy	547	TATCGGAGGTGGAGCGGAGGATCTCTGGCGCGCTGGCAGTGGAGGCGCTGGTAGGACT	606		
Db	590	TATCGGAGGTGGAGCGGAGGATCTCTGGCGCGCTGGCAGTGGAGGCGCTGGTAGGACT	649		
Qy	607	CGCAGGTGCTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCAGGTCTATGTACACAGCC	666		
Db	650	CGCAGGTGCTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCAGGTCTATGTACACAGCC	709		
Qy	667	TTCAAGCGCTCGGAGTTTCGGTGGACTACGTGCGAGGTCTGTGACTCGGACACAGGTTG	726		
Db	710	TTCAAGCGCTCGGAGTTTCGGTGGACTACGTGCGAGGTCTGTGACTCGGACACAGGTTG	769		
Qy	727	GACCCCATGCACTGCTGGAGCTCGTGGCGGTACTGGACGAGACACCCCGGTAGGGGCT	786		
Db	770	GACCCCATGCACTGCTGGAGCTCGTGGCGGTACTGGACGAGACACCCCGGTAGGGGCT	829		
Qy	787	GTTTGTGGGACGTGCGGATCTTTAAACCTCTCTGGACTCTTGGGTTCAGTCTCTTAAGCAGC	846		
Db	830	GTTTGTGGGATGTGCGGATCTTTAAACCTCTCTGGACTCTTGGGTTCAGTCTCTTAAGCAGC	889		
Qy	847	CTCGGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCACTGTGTA	906		
Db	890	CTCGGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCACTGTGTA	949		
Qy	907	TCTTGATCAGCGCTCTCTAGGCCCTATATAGGAATAACCTCTTGGACGAGTTTCTTGAG	966		
Db	950	TCTTGATCAGCGCTCTCTAGGCCCTATATAGGAATAACCTCTTGGACGAGTTTCTTGAG	1009		
Qy	967	GCCTGTGTACAAACAGAGTTCTCTGGGTACCACCTGTACTTTTGGGATACCGGCACCTC	1026		
Db	1010	GCCTGTGTACAAACAGAGTTCTCTGGGTACCACCTGTACTTTTGGGATACCGGCACCTC	1069		
Qy	1027	ACCAACCGCATGTCTAGCATGGGTTATGCTTACCAAGTA	1064		
Db	1070	ACCAACCGCATGTCTAGCATGGGTTATGCTTACCAAGTA	1107		

RESULT	5
LOCUS	ARI137038
DEFINITION	Sequence 1 from patent US 6162908.
ACCESSION	ARI137038
VERSION	ARI137038.1 GI:14478288
KEYWORDS	. Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified. 1 (bases 1 to 2117) Itano,N. and Kimata,K. Polypeptide of human-origin hyaluronate synthetase and DNA encoding the same
JOURNAL	Patent: US 6162908-A 1 19-DEC-2000;
FEATURES	Location/Qualifiers . . 2117 /organism="unknown" /mol_type="unassigned DNA"
ORIGIN	
Query Match	98.7%; Score 1051.6; DB 6; Length 2117;
Best Local Similarity	99.6%; Pred. No. 1.3e-134;

RESULT 6	AY463695	2109 bp	mrna	linear	PRI 07-DEC-2003
LOCUS	Papio anubis hyaluronan synthase (HAS1)				complete cds.
DEFINITION	AY463695				
ACCESSION	AY463695.1	GI:38607341			
VERSION					
KEYWORDS	Papio anubis (olive baboon)				
SOURCE	Papio anubis				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Papio.				
REFERENCE	1 (bases 1 to 2109)				
AUTHORS	Martinez-Duncker, I., Oriol, R. and Mollicone, R.				
TITLE	Evolution of the hyaluronan, nodulation c, chitin and cellulose synthases: a superfamily of cell-wall associated carbohydrate polymerizing enzymes				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 2109)				
AUTHORS	Martinez-Duncker, I., Oriol, R. and Mollicone, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-NOV-2003) Unite 504 Glycobiology Et Signalisation Cellulaire, Inserm, 16 Avenue Paul Vaillant Couturier, Villejuif 94807, France				
FEATURES	Location/Qualifiers				
source	1..2109				
	/organism="Papio anubis"				
	/mol_type="mrna"				
	/db_xref="taxon:9555"				
gene	1..2109				
	/gene="HAS1"				
CDS	36..1787				
	/gene="HAS1"				
	/note="cell wall associated protein; produces hyaluronic acid chains"				
	/codon_start=1				
	/product="hyaluronan synthase"				
	/protein_id="AAR25554.1"				
	/db_xref="GI:38607342"				
	/translation="MTQBDTPKTPAAPRCGSLARVLTIAPALLILGLMTWAYAAGV PLASRYGLLAFLGYAFSLAHLAQLSPFAYLEHRRVAARAAARGRIDAATARSV ALTI SAYQEDPAYRLVCLVSARALLYPRALRVLMVDGNRPEDLYWDMFRVFADE DPATYNYGNHQHPWPCAAVAGVAGYREVEADEPRLAVALVTRRCVCVAQRWG GKREVMYTFAPKALGSDVYVQVCSDTRLDPMALLQLVLDDEPRVAGVGVDRILN KLEVNSFLSLSLYVWAFNVBRACQSFHCVSCI SGPGLGRNNLLQOFLKAWYNKFL LGTHCTFGDDRHLTNEMLSGYATKYTSRSRCSYETPSSFLRWLSQOTRWSKSYFREW LYNALWHRHWRHMWYTEAVVSGLFPFPFVAATVLRFPAGRPWALLWVLLVCVQVALAK AAPAALWRCGLRWHLVSLYAPLTMCGLLPAKFLVLMQSGMTSGRRKLAAANYPL LPLIALWALLLLGLVRSVAHEARADWSGPRAAEAYVHLAAGAYGYVWVMTLYLV GVRRLCRRTTGGYRVQV"				
ORIGIN					
Query Match	92.6%	Score 986.2;	DB 9;	Length 2109;	
Best Local Similarity	95.6%	Pred. No. 1.1e-125;			
Matches 1031;	Conservative	0;	Mismatches 33;	Indels 15;	Gaps 1;
QY	1	ATGAGACAGCAGCAGCGCCCAAGCCACATCTCTGCAGCCCGCGCTGCTCCGCGCTGGCC	60		
Db	36	ATGACACAGCGGACACGCCCAAGCCACATCTCTGCAGCCCGCGCTGCTCCGCGCTGGCC	95		
QY	61	CGAGGGGTGTGACCATCGCTTTGCGCCTTCTCATCTTGGGCTCTATGACCTGGGCTTAC	120		
Db	96	CGAGGGGTGTGACCATCGCTTTGCGCCTTCTCATCTTGGGCTCTATGACCTGGGCTTAC	155		
QY	121	GCCGCCGGGGTCCGCTGGCTCCGATCGCTACGGCTCTCTGCGCTTCGCGCTCTACGGG	180		
Db	156	GCCGCCGGGGTCCGCTGGCTTCGATCGCTACGGCTCTCTGCGCTTCGCGCTCTACGGG	215		
QY	181	GCCTTCCTTTACGCGCACTGTGGCGCAGAGCCTCTTTCGCGTACCTGGAGCACCGCGCG	240		
Db	216	GCCTTCCTTCGCGCACTGTGTGGCGCAGAGCCTCTTCGCTACCTGGAGCATCGCGCG	275		

TITLE
Kannagi, R., Hamaguchi, M. and Kimata, K.
Selective expression and functional characteristics of three
Mammalian hyaluronan synthases in oncogenic malignant
transformation
J. Biol. Chem. 279 (18), 18679-18687 (2004)
14724275
2 (bases 1 to 2095)
Itano, N.
Direct Submission
Submitted (05-DEC-2002) Naoki Itano, Aichi Medical University,
Institute for Molecular Science of Medicine, Nagakute, Aichi, Aichi
480-1195, Japan (E-mail: itano@amugw.aichi-med-u.ac.jp,
Tel: 81-52-264-4811 (ex.2095), Fax: 81-561-63-3532)

FEATURES
Location/Qualifiers
1..2095
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1..2095
43..1794
/gene="HAS1"
/gene="HAS1"
/codon_start=1
/product="hyaluronan synthase 1"
/protein_id="BAC43730.1"
/db_xref="GI:26453349"
/translation="MRQMPKPESEAAACGSGLLARVLITTFALLILGLMTWYAAAGVP
LASDPVGLAFLYGLAPLSAHLVAQSLFAYLEHRRVTVAAERAFAPLDAATARSVA
LTIAYQEDPTVYQCLISARALLPRTFLRLVLMVVDGNRAEDLVYMDREPRVFADED
PATYVNDGNYHOPFAEAGAGGAEVREAEQDPLAVALVRLVRCVCAORWG
GHEVMYTAFAKAGSDVYQVDSDFDLDPMLLELVRVLEDEPRVGAAGDVRILN
PLDSVLSLSLYVAFVNERASQYFHCVSCI SGLPLRYNNLLQPFLEAWYQKF
LQTHCTFGDRHLTNRLMSMGYATKYSRSCYSETPSFLRWLQQTRWSKSYFEW
LYNALWHRHRRHMTLYEAVVSGLPFFVAAVTLVLFYAGRPWALLVLLLCVQVALAK
RAFAWLRCLRMVLLSLYAPLYMGLLPAKFLALVTVNQSGWGTSGRKKLAANTVPV
LPALWALLLLGLLRSAQAEVRADWSGPRAAEYHLLAAGASAYVAYVIMLTIIYW
GVRLLRRSSGGYRVQV"

ORIGIN

Query Match 74.3%; Score 791.4; DB 10; Length 2095;
Best Local Similarity 85.2%; Pred. No. 4,7e-99;
Matches 914; Conservative 0; Mismatches 141; Indels 18; Gaps 2;
QY 10 CAGGAGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCGGCGCTGCGCGAGGGTG 69
DB 49 CAGGACATGCCAAGCCCTCAGAGCAGCAGCTGCTGCTCGGCTGCGCAGCGGGTG 108
QY 70 CTGACCATCGCTTGCCTCTCATCTCTGGGCTCATGACCTGGGCTTACCGCGCGG 129
DB 109 CTCAGCATCACCTTGCCTCTCATCTCTGGGCTCATGACCTGGGCTTACCGCAGGA 168
QY 130 GTGCGCTGCTCGCATCGCTACGCGCTCTCGGCTCTGCGGCTCTACGGGCTTCTT 189
DB 169 GTACCTCTGGCTTGTATCCCTATGCGCTCTGCGCTTGGGCTCTATGGGCGTTCCTC 228
QY 190 TCAGCGCACCTGGTGGCGCAGAGCTCTTTCGGGTACTCTGAGCAGCGCGGCTGGCGG 249
DB 229 AGTGGCACCTAGTGGCAGAGCCCTTTCGCTTACCTGGAGCAGCGAGGCTGACCGTG 288
QY 250 GCGGCGCG-----GGGCGCGTGGATGACGACCGCGCGCGCATGTGGCG 294
DB 289 GCTGGCGCGCGCGCTTTTCGAAGGAGACCCCTGGATGCGGCGCTGCGCGCGTGGCA 348
QY 295 CTGACCATCTCGCTTACGAGGAGACCGCGCTACCTGCGCAGTGGCTGGCGCGCG 354
DB 349 CTCACCATCTCAGCTTACGAGGAGACCCCACTTACCTGCGCAGTGGCTTACCTCGCG 408
QY 355 CGCGCGCTCTGCTACCGCGCGCGGCTGCGGCTCTCATGCTGGTGGATGGCAACCGC 414
DB 409 CGCGCTTGTCTGTACCGCGCAGCGGCTGCGGCTGCTTATGGTTGTAGACGGCATCGC 468
QY 415 GCCGAGGACCTTACATGCTGCACATGTTTCGCGAGGCTTCTCGCTGACGAGGACCGCGC 474

DB 469 GCGGAGGATCTGTATCATGTTGGACATGTTCCGAGAAGTCTTTCGACATGAGGACCTGCC 528
QY 475 ACCTACGCTGTGGACGCGCAACTTACCACAGCCCTCGGAACCCGCG---GCGCGCGCGCG 531
DB 529 ACTTATGTTGGATGGCACTTACCATTAGCTTGGAGCCAGCTGAGGCGCGGGTGTCT 588
QY 532 GTGCGCGCGGAGCCTTATCGGAGGTGAGGCGGAGGATCTTGGCGGCTGGCAGTGGAG 591
DB 589 GTGGGTGAAGGTGCCCTACCGCGAGGTGAGGACCCCTGGCGCGCTGGCGGTAGAG 648
QY 592 GCGCTGTTGAGGACTCGCAGGTGCGTGTGCGTGGCGCGAGCGCTGGGCGGCAAGCGGAG 651
DB 649 GCGCTGTTGAGGACCGCAGGTGCGTGTGCGTGGCGCTGAGCGCTGGGCTGGCAAGCGGAA 708
QY 652 GTCATGTATACAGACCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCGAGTCTGTGAC 711
DB 709 GTCATGTATACAGGCTTTCAGGCACTGGGTGACTCCGTGGACTAGTGCAGGCTGTGAC 768
QY 712 TCGGACACAAGTTGGACCCCATGCGACTGTGAGCTGCTGCGGGTACTTGGACGAGGAC 771
DB 769 TCAGACACAAGTTAGACCCCATGCGACTGTGAGGCTTGTGCGAGTGTGATGAAGAC 828
QY 772 CCCGCGTAGGGCTGTTGGTGGGACCGTGGGATCCCTTAACCTCTGCACTCCCTGGTC 831
DB 829 CCCGCGTAGGGCTGTTGGTGGGAGAGATGTGAGATCTTAAACCTCTGCACTCTGGGTC 888
QY 832 AGCTTCTTAAGCAGCTCTCGATCTAGCTTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGC 891
DB 889 AGCTTCTTGGAGCGCTTGCATCTAGCTTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGC 948
QY 892 TACTTCCACTGTGTATCTCGCATCAGCGGTCTCTTAGGCTTATATAGGAATAACCTCTTG 951
DB 949 TACTTCTACTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008
QY 952 CAGCAGTCTTCTGAGCGCTGTGTAACAACAGAGTCTTCTGGGTACCCACTGTACTTCTGGG 1011
DB 1009 CAGCAGTCTTCTGAGCGCTGTGTAACAACAGAGTCTTCTGGGCGCCACTGTGCACTTGGG 1068
QY 1012 GATGACCGCGCACTTCAACCGCATGCTCAGCATGGTGTATGTACCAAGTA 1064
DB 1069 GATGACCGCGCACTTCAACCGCATGCTTAGCATGGTGTATGTACCAAGTA 1121
RESULT 8
E13681
LOCUS 2102 bp DNA linear PAT 27-APR-1998
DEFINITION DNA encoding novel mouse hyaluronate synthetase.
ACCESSION E13681
VERSION E13681.1 GI:3252450
KEYWORDS JP 1997224674-A/1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2102)
AUTHORS Itano, N. and Kimata, H.
TITLE POLYPEPTIDE OF NEW HYALURONIC ACID-SYNTHETIC ENZYME AND DNA CODING
JOURNAL THE SAME
Patent: JP 1997224674-A 1 02-SEP-1997;
SCIENCE & TECH AGENCY
OS Mus musculus (mouse)
PN JP 1997224674-A/1
PD 02-SEP-1997
PF 26-FEB-1996 JP 1996038336
PI ITANO NAKI, KIMATA HIROHARU
PC C12N1/09, C07K14/04, C07K14/47, C12N9/00, C12N1/21, C12N1/19; CC
(C12N9/00,
PC C12R1/91), (C12N9/00, C12R1/19), (C12N1/21, C12R1/19); CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers

```
PH source 1..2102
FT /organism='Mus musculus'
FT /cell_line='FM3A'
FT 5'UTR 1..48
FT CDS 49..1800
FT /product='Novel mouse hyaluronate synthetase'
FT 3'UTR 1801..2102.
FEATURES
    source
        1..2102
        /organism='Mus musculus'
        /mol_type='genomic DNA'
        /db_xref='taxon:10090'
ORIGIN
Query Match 73.9%; Score 786.6; DB 6; Length 2102;
Best Local Similarity 84.9%; Pred. No. 2.1e-98;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;
QY 10 CAGGAGCGCCCAAGCCCACTCTGAGCCGCGCGCTGCTCGGCTGGCCCGGAGGGTG 69
DB 55 CAGGACATGCCAAGCCCTCAGAGCAGCGCGTTCCTGCTCTGGCTTGGCCAGCGGAGCA 114
QY 70 CTGACCATCGCTTCGCCCTGCTCATCTCGGCTCATGACCTGGGCTTACGCGCGCGG 129
DB 115 CTCAGATCATCTTTGGCCCTGCTCATCTCGGCTCATGACCTGGGCTTACGCGCGAGC 174
QY 130 GTGCGCTGGCTTCGATCGCTACGCGCTCTGGGCTTGGGCTCTACGCGGCTTCTT 189
DB 175 GTTCTCTGGCTTCAGATCGCTATGGACTCTCTGGCTTGGGCTCTATGGGGCATTCCTC 234
QY 190 TCAGCGCACCTGGTGGCGAGAGCTCTTCGGTACCTGGAGCACCGGCGGGTGGCGG 249
DB 235 AGCGCACACCTAGTGGCACAGAGCTCTTCGTTACCTGGAGCACCGAAGGGTGGCAGC 294
QY 250 GCGGCGCG-----GGGCGCGCTGGATGACGACCCAGCCGCGCGAGTGTGGCG 294
DB 295 GTTGGCGGCGCTCTTGGCGAAGGGGCCCTTGGATGGCGCACTGACGACGAGTGGCA 354
QY 295 GTGACATCTCGCTACAGAGAGACCCCGCTTACCTGGCGCACTGCTGGCGTTCGCGC 354
DB 355 CTACCATCTCAGCCTACCAAGAGATCCCGCTTACCTGGCGCACTGCTGACCTTCGCG 414
QY 355 CGGCGCTGCTTACCGCGCGCGGCTGCGGCTCTCATGGTGGTGGATGGCAACGCG 414
DB 415 CGCGCTTGTCTACCGCGCACAGAGGTTACGCGTCTCATGGTGGTGGACGCGCAACGCG 474
QY 415 GCGGAGGACCTCATATGGTGACATGTTCCGCGAGGTTCTCGCTGACGAGGACCCCGCC 474
DB 475 GCTGAGGATCTGTATATGGTGACATGTTCCGAGAGTCTTCGCGATGAGGACCCCGCC 534
QY 475 ACGTACGTGTGGACGCGCAACTACCAAGAGCTTGGGAAACCGCGCGCG---CGGCGCG 531
DB 535 ACTTATGTGGGATGGCAACTTACCATCAGCCCTGGGAACCAAGCGGCTACGCGGCT 594
QY 532 GTGGCGCGGAGCTATCGGAGGTGAGGCGGAGGATCTTGGCGGCTGGCAGTGGAG 591
DB 595 GTGCGTGAAGTGTCCTACCGGAGGTGAGGCGGAGGACCCCGCGGCTTGGCGGTGGAG 654
QY 592 GCGTGGTGAGGACTCGCAGGTGCGTGGCGGCGAGCGCTGGGCGCGCAAGCGGAG 651
DB 655 GCGTGGTGAGNACACGACGATGCTGTGGCTGCTGCGGCTTGGGCGGCAAGCTGAG 714
QY 652 GTATGTACACAGCTTCAAGCGCTTCGAGATTCGGTGGACTACGTGCAAGTCTGTGAC 711
DB 715 GTATGTACACAGCTTCAAGGCACTGGGCGACTCCGTGGACTACGTGCAAGTCTGTGAC 774
QY 712 TCGRACAGAGTTGGACCCCATGCACTGTGGAGCTCGTGGGCTACTGACGAGGAC 771
DB 775 TCAGACACAGAGCTAGACCCCATGCACTGTGTGGAGCTGTGTGGAGTGTGTGATGAAGAC 834
QY 772 CCGCGGGTAGGGGCTGTGTGGGAGCGTGGGATCTTAAACCTCTCGACTCTCTGGGTC 831
```

```
DB 835 CCGCGGGTAGGGCTGTTGGAGGGGATGTGAGGATCTTAAACCTCTCGACTCTGGGTC 894
QY 832 AGCTTCTTAACAGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGACAGC 891
DB 895 AGCTTCTTGGAGCAGCTTTCGATACTCTGGGTAGCCTTCAATGTGGAAACGAGCTTGTGACAGC 954
QY 892 TACTTCCACTGTGTATCTCTGATCAGCGGTCTCTAGGCTTATATAGGAATAACCTCTTG 951
DB 955 TACTTCCACTGTGTGTCTCTGATCAGTGTCTCTGGGTCTATACAGAAACAATCTCTG 1014
QY 952 CAGCAGTTTCTTGGAGGCTGTGTACAAACAGAAAGTTCTTGGGTACCCACTGTTTGGG 1011
DB 1015 CAGCAGTTTCTTGGAGGCTGTGTACAAACAGAAAGTTCTTGGGACCCACTGTCACATTTGGG 1074
QY 1012 GATGACCGGACCTTACCAACCGCATGCTAGCATGGGTTATGTACCAAGTA 1064
DB 1075 GATGACAGGCACCTTACCAACCGAATGCTTAGCATGGGCTATGTACCAAGTA 1127
RESULT 9
E28454 LOCUS 2102 bp DNA linear PAT 18-JUN-2001
DEFINITION Hyaluronate synthase promoter DNA.
ACCESSION E28454
VERSION E28454.1 GI:13018346
KEYWORDS JP 1999196875-A/2.
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE 1 (bases 1 to 2102)
AUTHORS Yoichi,Y., Naoki,I. and Koji,K.
TITLE Hyaluronate synthase promoter DNA
JOURNAL Patent: JP 1999196875-A 2 27-JUL-1999;
SEIKAGAKU KOGYO CO LTD
COMMENT OS Mus sp. (mouse)
PD JP 1999196875-A/2
PD 27-JUL-1999
PF 14-JAN-1998 JP 1998006191
PR YOICHI YAMADA,NAOKI ITANO,KOJI KIMATA
PC C12N15/09,C12N9/00,C12Q1/68//((C12N15/09,C12R1:91),C12N15/00,
(C12N15/00,C12R1:91)
PC
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS
    Location/Qualifiers
        1..2102
        /organism="Mus sp."
        /mol_type="genomic DNA"
        /db_xref="taxon:10095"
ORIGIN
Query Match 73.9%; Score 786.6; DB 6; Length 2102;
Best Local Similarity 84.9%; Pred. No. 2.1e-98;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;
QY 10 CAGGAGCGCCCAAGCCCACTCTCTGAGCCGCGCGCTGCTCGGCTGGCCCGGAGGGTG 69
DB 55 CAGGACATGCCAAGCCCTCAGAGGAGCGCGTTCCTGCTCTGGCTTGGCCAGCGGAGCA 114
QY 70 CTGACCATCGCTTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTACGCGCGCGG 129
DB 115 CTCAGATCATCTTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTACGCGCGAGC 174
QY 130 GTGCGCTGGCTTCGATCGCTACGCGCTCTGGGCTTGGGCTCTACGCGGCTTCTT 189
DB 175 GTTCTCTGGCTTCAGATCGCTATGGACTCTCTGGCTTGGGCTCTATGGGGCATTCCTC 234
QY 190 TCAGCGCACCTGGTGGCGAGAGCTCTTTCGCTTACCTGGAGCACCGGCGGGTGGCGG 249
DB 235 AGCGCACACCTAGTGGCACAGAGCTCTTCGTTACCTGGAGCACCGAAGGGTGGCAGC 294
```


Db 775 TCAGACAAAGACTAGACCCCATGGCACTGCTGGAGCTTGTGCGAGTCTTGGATGAAGAC 834
Qy 772 CCCCAGGAGAGGCTGTTGGTGGGACGTGCGGATCCCTTAACCCCTCTGGAATCTCTGGGTC 831
Db 835 CCCCAGGAGAGGCTGTTGGAGGGGATGTGAGGATCCCTTAACCCCTCTGGAATCTCTGGGTC 894
Qy 832 AGCTTCTTAAGCAGCTCGGATACCTGAGTACCTGAGTACCTCAATGTGGAGCGGCTTGTGAGGC 891
Db 895 AGCTTCTTAAGCAGCTCGGATACCTGAGTACCTGAGTACCTCAATGTGGAGCGGCTTGTGAGGC 954
Qy 892 TACTTCCACTGTGTATCTCTGCATCAGCGGTCTCTAGGCCTATATAGGAATAAACCTCTTG 951
Db 955 TACTTCCACTGTGTATCTCTGCATCAGTGTCTCTGGGTCTATACAGAAAATCTCTCTG 1014
Qy 952 CAGCAGTTCCTTGGAGGCTGTGTACAAACAGAAAGTTCCTGGGTACCCACTGTGACTTTTGGG 1011
Db 1015 CAGCAGTTCCTTGGAGGCTGTGTACAAACAGAAAGTTCCTGGGTACCCACTGTGACATTTGGG 1074
Qy 1012 GATGACCGGACCTCACCACCGCATGCTCAGCATGGGTATGCTACCAAGTA 1064
Db 1075 GATGACCGGACCTCACCACCGCATGCTTACGATGGGCTATGCTACCAAGTA 1127

RESULT 11

E34326 E34326 2102 bp DNA linear PAT 31-JAN-2002
LOCUS DNA for gene targeting hyaluronic acid synthase gene.

DEFINITION E34326

VERSION E34326.1 GI:18624311

KEYWORDS JP 2000116382-A/1.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE Itano,N. and Kimata,K.

JOURNAL DNA for gene targeting hyaluronic acid synthase gene

COMMENT Patent: JP 2000116382-A 1 25-APR-2000;

SEIKAGAKU KOGYO CO LTD

OS Mus musculus (mouse)

PN JP 2000116382-A/1

PD 25-APR-2000

PF 13-OCT-1998 JP 1998291201

PR NAOKI ITANO,KOJI KIMATA

PC C12N15/09,C12N5/10,G01N33/50//C12N9/00,C12Q1/68,(C12N15/09,PC

C12R1:91),

PC (C12N5/10,C12R1:91),C12N5/00,C12N5/00,(C12N15/00,C12R1:91),

CC (C12N5/00,C12R1:91)

FT CDS Location/Qualifiers

1..2102

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

FEATURES

source

ORIGIN

Query Match 73.9%; Score 786.6; DB 6; Length 2102;

Best Local Similarity 84.9%; Pred. No. 2.1e-98;

Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;

Qy 10 CAGACGGCGCCCAAGCCCACTCTCGACGCGCGCTGCTCGGCTTGGCGGAGGGTG 69

Db 55 CAGGACATGCCAAGCCCTCAGAGCAGCGGTGCTGCTTGGCTTGGCCAGCGGACGA 114

Qy 70 CTGACATCGGCTTGGCGCTGCTATCTTGGGCTCATGACCTGGGCTTACGCGCGGG 129

Db 115 CTCAGATCATCTTTGGCGCTGCTATCTTGGGCTCATGACCTGGGCTTACGCGCAGGC 174

Qy 130 GTGCGCTGGCTCGGATCGCTACGCGCTCTGGGCTTGGGCTTACGCGGCTTCTT 189

Db 175 GTTCTCTGGCTTCAGATCGCTATGAGACTCTGGCTTTTGGCCTCTATGCGGCATTCCTC 234
Qy 190 TCAGGCACTGTGTGGGCGCAGAGCCTCTTTCGCTACCTGAGACACCGCGCGGTGGCGGC 249
Db 235 AGCGCACACCTAGTGGCACAGAGCCTCTTTCGCTTACCTTGGAGCACCGAAGGTTGGCAGC 294
Qy 250 GCGCGCGG-----GGGCGCGCTGTGATGACGACCGCGCGCAGTGTGGCG 294
Db 295 GCTCGGCGGCGCTCTTGGCGAAGGGGCGCTGTGATGCGGCCATGTGACGAGCGTGGCA 354
Qy 295 CTGACCATCTCCGCTTACCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354
Db 355 CTACCATCTCAGCTTACCGAG 414
Qy 355 CGCGCCCTGTGTATACCGCGCGCGGCTGGCGCTCTCATGTGTGGTGGATGGCAACGCG 414
Db 415 CGCGCCCTGTGTATACCGCGCACACGAGGTTCATGCGTGTCTCATGTGTGGTGGAGCGCAACGCG 474
Qy 415 GCGGAGGACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGAGGAGAGAGAGAGAG 474
Db 475 GCTGAGGATCTGTATGTTGAGGAGATGTTCCGAGAGGTCTTTCGCGGATGAGAGAGAGAGAG 534
Qy 475 ACGTACGTGTGGGACGGCACTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 531
Db 535 ACTTATGTGGGATGGCACTACATCAGCCCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594
Qy 532 GTGGGCGCGAG 591
Db 595 GTCGGTGAAGTGTCTACCGGAGGTGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
Qy 592 GCGCTGTGTAGGAGCTCGCAGGTGTGTGCGTGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAG 651
Db 655 GCGCTGTGTAG 714
Qy 652 GTCATGTACAG 711
Db 715 GTCATGTACAG 774
Qy 712 TCGGACACAGAGTGTGGAG 771
Db 775 TCAGACACAGAGACTAGACCCCATGGCCTGTGCGAGGTGTGCGAGTGTGGATGAAGAC 834
Qy 772 CCGCGGTAGGGGCTGTGTGGGAGCTGCGGATCCTTAACTCTGGAGCTCTCTGGGTC 831
Db 835 CCGCGGTAGGGGCTGTGTGGGAG 894
Qy 832 AGCTTCTTAAGCAG 891
Db 895 AGCTTCTTAAGCAG 954
Qy 892 TACTTCCACTGTGTATCTTGCATCAGCGGTCTCTAGGCTTATATAGGAATAACCTCTTG 951
Db 955 TACTTCCACTGTGTATCTTGCATCAGTGTCTCTGGGCTCTATACAGAAAATCTCTCTG 1014
Qy 952 CAGCAGTTCCTTGGAGGCTGTGTACAAACAGAAAGTTCCTGGGTACCCACTGTGACTTTTGGG 1011
Db 1015 CAGCAGTTCCTTGGAGGCTGTGTACAAACAGAAAGTTCCTGGGACCCACTGTGACATTTGGG 1074
Qy 1012 GATGACCGGACCTCACCACCGCATGCTCAGCATGGGTATGCTACCAAGTA 1064
Db 1075 GATGACAGGACCTCACCACCGCATGCTTACGATGGGCTATGCTACCAAGTA 1127

RESULT 12

LOCUS

MUSHAS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

MUSHAS 2102 bp mRNA linear ROD 06-FEB-1999
Mus musculus mRNA for hyaluronan synthase, complete cds.
D82964.1 GI:1339939
hyaluronan synthase.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```
Db 146930 GGGGTGCGCTGGCTCCGATCGCTACGCGCTCTCGGCTTCTACGGCGGCTTC 146871
QY 187 CTTTCAGCGCACTGGTGGCGCAGAGCTCTTCGGTACCTGGAGCACCGCGGGTGGCG 246
Db 146870 CTTTCAGCGCACTGGTGGCGCAGAGCTCTTCGGTACCTGGAGCACCGCGGGTGGCG 146811
QY 247 GCGGCGGCGCGGGGCGGCTGGATGACACCGCGCGCAGTGTGGCGCTGACCATCTCC 306
Db 146810 GCGGCGGCGCGGGGCGGCTGGATGACACCGCGCGCAGTGTGGCGCTGACCATCTCC 146751
QY 307 GCCTACCAAGGAGACCCCGCTACTGCGCCAGTGCCTGGCGTCCGCCCGCGCCCTGCTG 366
Db 146750 GCCTACCAAGGAGACCCCGCTACTGCGCCAGTGCCTGGCGTCCGCCCGCGCCCTGCTG 146691
QY 367 TACCGCGCGCGGCTGGCGTCTCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 426
Db 146690 TACCGCGCGCGGCTGGCGTCTCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 146631
QY 427 TACATGTTGCAATGTTTCCGCGAGGTCTTCTGTCACGAGGACCCCGCCACGTCAGTGTGG 486
Db 146630 TACATGTTGCAATGTTTCCGCGAGGTCTTCTGTCACGAGGACCCCGCCACGTCAGTGTGG 146571
QY 487 GACGGCACTACCAACGACCCCTGGGAAACCCCGCGCGCGCGCGTGGCGCGCGGAGCC 546
Db 146570 GACGGCACTACCAACGACCCCTGGGAAACCCCGCGCGCGCGCGTGGCGCGCGGAGCC 146511
QY 547 TATCGGAGGTGGAGCGGAGATCTTGGCGGCTGGCAGTGGAGGCGCTGTTGAGGACT 606
Db 146510 TATCGGAGGTGGAGCGGAGATCTTGGCGGCTGGCAGTGGAGGCGCTGTTGAGGACT 146451
QY 607 GCGAGGTGCGTGTGGTGGCGCAGCGCTGGGGCGCGCAAGCCGCGAGGTCTATGTACACAGCC 666
Db 146450 GCGAGGTGCGTGTGGTGGCGCAGCGCTGGGGCGCGCAAGCCGCGAGGTCTATGTACACAGCC 146391
QY 667 TTCAAGCGCTCGGAGATTCCGTGAGCTACGTGACAGTCTCT 708
Db 146390 TTCAAGCGCTCGGAGATTCCGTGAGCTACGTGAGTGTGAGTGTGAGT 146349

RESULT 15
AC130783/c
LOCUS
DEFINITION Pan troglodytes clone CH251-426A12, WORKING DRAFT SEQUENCE, 8 linear HTG 19-NOV-2002
ACCESSION AC130783
VERSION AC130783.2 GI:25100968
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (Bases 1 to 193986)
Akhtar,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaca,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.I., Maduro,V.B.,
Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Duque,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
JOURNAL Unpublished
REFERENCE 2 (Bases 1 to 193986)
Green,E.D.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Grosvmont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (Bases 1 to 193986)
Green,E.D.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2002) NIH Intramural Sequencing Center, 8717
Grosvmont Circle, Gaithersburg, MD 20877, USA
```

COMMENT

Grovemont Circle, Gaithersburg, MD 20877, USA
On Nov 19, 2002 this sequence version replaced gi:22218453.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: dxd
Center clone name: 426A12

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 191764 bases at least Q40
Consensus quality: 192510 bases at least Q30
Consensus quality: 192998 bases at least Q20
Insert size: 184000; agarose-fp
Quality coverage: 11.49x in Q20 bases; agarose-fp
Quality coverage: 10.94x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.
* 11512: contig of 11512 bp in length
* 11513: gap of unknown length
* 11613: contig of 47876 bp in length
* 59489: gap of unknown length
* 59589: contig of 43678 bp in length
* 103267: gap of unknown length
* 103367: contig of 24519 bp in length
* 127886: gap of unknown length
* 127986: contig of 1707 bp in length
* 129692: gap of unknown length
* 129793: contig of 1285 bp in length
* 131078: gap of unknown length
* 131178: contig of 43486 bp in length
* 174664: gap of unknown length
* 174764: contig of 19223 bp in length.

FEATURES

source

1. 193986
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-426A12"
/clone_lib="CH251"

misc_feature

1. 11512
/note="assembly_fragment
clone_end:SP6
vector_side:left"

misc_feature

1. .76
/note="clone overlaps with GenBank Accession Number
AC132069 clone CH251-132A6 (center project name dxc)"

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 01:47:21 ; Search time 692.135 Seconds
(without alignments)
9108.809 Million cell updates/sec

Title: US-10-672-399-7
Perfect score: 1065
Sequence: 1 atgagacagcaggacgccc.....tgggttatgtaccacagtaa 1065

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	99.9	1737	ADL13691	Adl13691 Osteoarth
2	1064	99.9	2088	ADL13690	Adl13690 Osteoarth
3	1064	99.9	2116	AAT99541	Aat99541 Human hya
4	1062.4	99.8	1737	ADL13692	Adl13692 Osteoarth
5	1062.4	99.8	2087	ADL13694	Adl13694 Osteoarth
6	1042	97.8	2117	AAT96713	Aat96713 Human hya
7	786.6	73.9	1752	ABZ76734	Abz76734 Mouse hya
8	786.6	73.9	1752	AAZ59442	Aaz59442 Mouse hya
9	786.6	73.9	2102	AAT91655	Aat91655 Mouse hya
10	786.6	73.9	2102	AAZ10862	Aaz10862 Hyaluroa
11	786.6	73.9	2102	AAZ88199	Aaz88199 Mouse hya
12	786.6	73.9	2102	AAZ39987	Aaz39987 Murine HA
c 13	697.2	65.5	231222	ADL13693	Adl13693 Osteoarth
14	422	39.6	662	ABQ27384	Abq27384 Oligonuc
c 15	422	39.6	662	ABQ27385	Abq27385 Oligonuc
16	421	39.5	490	ACH39998	Ach39998 Human foe
c 17	393.8	37.0	662	ABQ27382	Abq27382 Oligonuc
18	393.8	37.0	662	ABQ27383	Abq27383 Oligonuc
19	366	34.4	1783	ADC49212	Adc49212 Rabbit hy
20	353	33.1	1662	ACC51023	Acc51023 Human bla

21	353	33.1	1662	8	ABX76415	Abx76415 Lung canc
22	353	33.1	1662	11	ADN39051	Adn39051 Cancer/an
23	353	33.1	4049	8	ABX76214	Abx76214 Lung canc
24	353	33.1	4049	12	ADN05115	Adn05115 Antipeori
25	353	33.1	4220	12	ADJ74905	Adj74905 Marker ge
26	323.8	30.4	1767	10	ADD93927	Add93927 Xenopus l
27	311.6	29.3	1665	2	AAV18822	Aav18822 Murine hy
28	311.6	29.3	1665	3	AAZ88201	Aaz88201 Mouse hya
29	311.6	29.3	1665	8	ABZ76736	Abz76736 Mouse hya
30	311.6	29.3	1665	10	AAZ59444	Aaz59444 Mouse hya
31	311.6	29.3	5919	12	ADJ75814	Adj75814 Marker ge
c 32	246.8	23.2	1653	6	ABL41013	Ab141013 Murine hy
33	246.8	23.2	1659	8	ABZ76735	Abz76735 Mouse hya
34	246.8	23.2	1659	10	AAZ59443	Aaz59443 Mouse hya
35	246.8	23.2	2947	2	AAV18821	Aav18821 Murine hy
36	246.8	23.2	2948	6	ABI99697	Abi99697 Mouse lsc
37	246.8	23.2	4194	9	ADA49685	Ada49685 DNA encod
38	246.8	23.2	4194	10	ADC59298	Adc59298 Mouse HAS
39	233.2	21.9	2890	3	AAZ88200	Aaz88200 Mouse hya
40	221	20.8	1659	10	ADL13696	Adl13696 Osteoarth
41	221	20.8	3003	6	ABK84373	Abk84373 Human CDN
42	221	20.8	3003	9	ADA49683	Ada49683 DNA encod
43	221	20.8	3003	10	ADC59296	Adc59296 Human HAS
44	221	20.8	3003	10	ADD12718	Add12718 Human CDN
45	221	20.8	3003	10	ADL13695	Adl13695 Osteoarth

ALIGNMENTS

RESULT 1

ADL13691
ID ADL13691 standard; DNA; 1737 BP.
XX
AC ADL13691;
XX
DT 06-MAY-2004 (first entry)
XX
DE Osteoarthritis-associated polymorphic nucleotide #223.
XX
KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FN WO2003054166-A2.
XX
PD 03-JUL-2003.
XX
PF 19-DEC-2002; 2002WO-US041225.
XX
PR 20-DEC-2001; 2001US-0342603P.
XX
(INCY-) INCYTE GENOMICS INC.
XX
Jones KA, Schafer A;
XX
WPI; 2003-559141/52.
XX
Determining susceptibility of an individual to joint space narrowing,
XX
osteophyte development and/or joint pain comprises identifying whether
XX
the individual has at least one polymorphism in a polynucleotide encoding
XX
a protein.
XX
Disclosure; SEQ ID NO 223; 297pp; English.
XX
The invention relates to a method of determining susceptibility of an
XX
individual to joint space narrowing and/or osteophyte development and/or
XX
joint pain comprising identifying whether the individual has at least one
XX
polymorphism in a polynucleotide encoding at least one of the protein
XX
listed in the specification. The methods, composition and agent are
XX
useful for modulating the susceptibility of an individual to joint space

CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).

XX SQ Sequence 1737 BP; 232 A; 571 C; 603 G; 331 T; 0 U; 0 Other;

Query Match 99.9%; Score 1064; DB 10; Length 1737;

Best Local Similarity 100.0%; Pred. No. 6.7e-180; Indels 0; Gaps 0;

Matches 1064; Conservative 0; Mismatches 0;

QY 1 ATGAGACAGCAGGAGCGCCCAAGCCCACTCCTGCGAGCCCGCGCTGCTCGGCGCTGGCC 60
DB 1 ATGAGACAGCAGGAGCGCCCAAGCCCACTCCTGCGAGCCCGCGCTGCTCGGCGCTGGCC 60
QY 61 CGGAGGGTGTGACCATCGCTTCGCGCTTCATCTCGGCGCTCATGACCTGGGCGCTTAC 120
DB 61 CGGAGGGTGTGACCATCGCTTCGCGCTTCATCTCGGCGCTCATGACCTGGGCGCTTAC 120
QY 121 GCGCGCGGGTGCCTGCGCTCCGATCGCTACGCGCTTCGCGCTTCGCGCTTCGCGCTTCACGG 180
DB 121 GCGCGCGGGTGCCTGCGCTCCGATCGCTACGCGCTTCGCGCTTCGCGCTTCGCGCTTCACGG 180
QY 181 GCCTTCCTTTTCAGCGCACCTGCTGGCGCAGAGCCTTTCGCGCTTCGCGCAGCAGCCGCGG 240
DB 181 GCCTTCCTTTTCAGCGCACCTGCTGGCGCAGAGCCTTTCGCGCTTCGCGCAGCAGCCGCGG 240
QY 241 GTGGCGGCGGCGCGCGCGCGCGCTGTGATGAGCAGCAGCAGCGCGCAGTGTGGCGCTGAC 300
DB 241 GTGGCGGCGGCGCGCGCGCGCGCTGTGATGAGCAGCAGCAGCGCGCAGTGTGGCGCTGAC 300
QY 301 ATCTCGGCTACACAGAGAGACCCCGCTACCTGCGCGCAGTGCCTGCGCGCGCGCGCC 360
DB 301 ATCTCGGCTACACAGAGAGACCCCGCTACCTGCGCGCAGTGCCTGCGCGCGCGCGCC 360
QY 361 CTGCTGTATCCCGCGCGCGCGCTGCGCGCTCCTCATGTGTGGATGGCAACCGCGCGCGAG 420
DB 361 CTGCTGTATCCCGCGCGCGCGCTGCGCGCTCCTCATGTGTGGATGGCAACCGCGCGCGAG 420
QY 421 GACCTCTACATGTCGATGTCGCGAGGTCCTGCTGAGCAGAGACCCCGCGCAGTAC 480
DB 421 GACCTCTACATGTCGATGTCGCGAGGTCCTGCTGAGCAGAGACCCCGCGCAGTAC 480
QY 481 GTGTGGGACGCACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGCGCGCGCGCC 540
DB 481 GTGTGGGACGCACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGCGCGCGCGCC 540
QY 541 GGAGCGCTATCGGAGGTGAGCGGAGGATCCTGCGCGCGCTGGCAGTGGAGCGCGCTGGTG 600
DB 541 GGAGCGCTATCGGAGGTGAGCGGAGGATCCTGCGCGCGCTGGCAGTGGAGCGCGCTGGTG 600
QY 601 AGGACTCGAGGTGCTGTGTCGCGCAGCGCTGGGCGGCGCAACCGCGAGGTCTATGTAC 660
DB 601 AGGACTCGAGGTGCTGTGTCGCGCAGCGCTGGGCGGCGCAACCGCGAGGTCTATGTAC 660
QY 661 ACAGCGCTTCAGGCGCTCGGAGTTTCGGTGGACTACGTGCGAGTCTGTGACTCGGACACA 720
DB 661 ACAGCGCTTCAGGCGCTCGGAGTTTCGGTGGACTACGTGCGAGTCTGTGACTCGGACACA 720
QY 721 AGGTTGAGCCCATGCGCATGCTGTGAGTCTGTCGGGCTACTGGACGAGACCCCGGCTA 780
DB 721 AGGTTGAGCCCATGCGCATGCTGTGAGTCTGTCGGGCTACTGGACGAGACCCCGGCTA 780
QY 781 GGGGCTGTGTGGGGAAGTCGGAGTCTTTAACCTCTGAGACTCTCGGGTCACTTCTTA 840
DB 781 GGGGCTGTGTGGGGAAGTCGGAGTCTTTAACCTCTGAGACTCTCGGGTCACTTCTTA 840
QY 841 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGTCAGAGCTACTTCCAC 900

DB 841 AGCAGCCTGCGATACTGGTACGCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCAC 900
QY 901 TGTGTATCTGTCATCAGCGGTCTCTAGGCTATATAGGAATAACCTCTTGAGCAGTTT 960
DB 901 TGTGTATCTGTCATCAGCGGTCTCTAGGCTATATAGGAATAACCTCTTGAGCAGTTT 960
QY 961 CTTGAGGCGCTGGTACACCAAGAGTTCTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
DB 961 CTTGAGGCGCTGGTACACCAAGAGTTCTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTGTATGTACCAAGTA 1064
DB 1021 CACCTCACCAACCGCATGCTCAGCATGGTGTATGTACCAAGTA 1064

RESULT 2

ADL13690

ID ADL13690 standard; DNA; 2088 BP.

AC ADL13690;

XX 06-MAY-2004 (first entry)

XX Osteoarthritis-associated polymorphic nucleotide #222.

XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;

XX joint space narrowing; osteophyte development; joint pain;

XX osteoarthritis; SNP; single nucleotide polymorphism.

XX Homo sapiens.

XX WO2003054166-A2.

XX 03-JUL-2003.

XX 19-DEC-2002; 2002WO-US041225.

XX 20-DEC-2001; 2001US-0342603P.

XX (INCY-) INCYTE GENOMICS INC.

XX Jones KA, Schafer A;

XX WPI; 2003-559141/52.

XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.

XX Disclosure; SEQ ID NO 222; 297pp; English.

XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polynucleotide encoding a protein listed in the specification. (Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences).

XX Sequence 2088 BP; 302 A; 652 C; 711 G; 423 T; 0 U; 0 Other;

Query Match

Best Local Similarity 99.9%; Score 1064; DB 10; Length 2088;

100.0%; Pred. No. 6.7e-180;

Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGAGACAGCAGGCGCGCCCAAGCCACTCTCTGAGCGCGCGCTGCTCCGGCTGGCC	60						
Db	36	ATGAGACAGCAGGCGCGCCCAAGCCACTCTCTGAGCGCGCGCTGCTCCGGCTGGCC	95						
QY	61	CGGAGGCTGTGACCATCGCTTGGCTTGTCTATCTTGGGCTCATGACCTGGGCTTAC	120						
Db	96	CGGAGGCTGTGACCATCGCTTGGCTTGTCTATCTTGGGCTCATGACCTGGGCTTAC	155						
QY	121	CGCGCGGCGGCTCGCTCCGATCGTACGGCTCTTGGCTTGGCTTACCGG	180						
Db	156	CGCGCGGCGGCTCGCTCCGATCGTACGGCTCTTGGCTTGGCTTACCGG	215						
QY	181	GCCTTCCTTTACGCGACCTGTGTGCGCAGAGCCTCTTTCGCTACCTTGAGCAGCGGG	240						
Db	216	GCCTTCCTTTACGCGACCTGTGTGCGCAGAGCCTCTTTCGCTACCTTGAGCAGCGGG	275						
QY	241	GTGGCGGCGGCGCGGCGGCGCTGTGATGAGCAGCAGCGCGCAGTGTGGCGCTGACC	300						
Db	276	GTGGCGGCGGCGCGGCGGCGCTGTGATGAGCAGCAGCGCGCAGTGTGGCGCTGACC	335						
QY	301	ATCTCCGCTACACAGAGAGACCCCGCTACCTGCGCAGTGTGCTGCGCGCGCC	360						
Db	336	ATCTCCGCTACACAGAGAGACCCCGCTACCTGCGCAGTGTGCTGCGCGCGCC	395						
QY	361	CTGCTGTATACCCGCGCGCGCTGCGCTCTCATGTGGTGGATGGCAACCGCGCCGAG	420						
Db	396	CTGCTGTATACCCGCGCGCGCTGCGCTCTCATGTGGTGGATGGCAACCGCGCCGAG	455						
QY	421	GACCTCTATAGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC	480						
Db	456	GACCTCTATAGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC	515						
QY	481	GTGTGGAGCGCAATACACAGACCTTGGGAACCCGCGCGCGCGCGCTGGCGCC	540						
Db	516	GTGTGGAGCGCAATACACAGACCTTGGGAACCCGCGCGCGCGCGCTGGCGCC	575						
QY	541	GAGGCTATCGGAGGTGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGTG	600						
Db	576	GAGGCTATCGGAGGTGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGTG	635						
QY	601	AGGACTCGAGGTGCTGTGCTGCGCAGCGCTGCGCGCGCAAGCGAGGTCTATGTAC	660						
Db	636	AGGACTCGAGGTGCTGTGCTGCGCAGCGCTGCGCGCGCAAGCGAGGTCTATGTAC	695						
QY	661	ACAGGCTTCAAGCCCTCGAGATTCGCTGAGTACGTGAGTCTGTGACTCGGACACA	720						
Db	696	ACAGGCTTCAAGCCCTCGAGATTCGCTGAGTACGTGAGTCTGTGACTCGGACACA	755						
QY	721	AGTTTGGACCCCATGCTGCTGAGCTCGTGGGCTACTGGACGAGGACCCCGGTA	780						
Db	756	AGTTTGGACCCCATGCTGCTGAGCTCGTGGGCTACTGGACGAGGACCCCGGTA	815						
QY	781	GGGCTGTGTGGGACGTGCGGATCTTAACTCTGACTCTGGCTCAGCTTCTTA	840						
Db	816	GGGCTGTGTGGGACGTGCGGATCTTAACTCTGACTCTGGCTCAGCTTCTTA	875						
QY	841	AGCAGCTCGGATATCTGGGTACCTTCAATGTGGAGCGGCTTGTGACAGTACTTCCAC	900						
Db	876	AGCAGCTCGGATATCTGGGTACCTTCAATGTGGAGCGGCTTGTGACAGTACTTCCAC	935						
QY	901	TGTGTATCTGATCAGCGCTCTCTAGCGCTATATAGGATTAACCTCTTGCAGAGTTT	960						
Db	936	TGTGTATCTGATCAGCGCTCTCTAGCGCTATATAGGATTAACCTCTTGCAGAGTTT	995						
QY	961	CTTGAGGCTGTGTACAAACAGAGTTCCTGGGTACCCACTGTACTTTTGGGATCACCGG	1020						
Db	996	CTTGAGGCTGTGTACAAACAGAGTTCCTGGGTACCCACTGTACTTTTGGGATCACCGG	1055						
QY	1021	CACCTCACACCGCATGCTAGCATGGTGTATGTACCAAGTA	1064						
Db	1056	CACCTCACACCGCATGCTAGCATGGTGTATGTACCAAGTA	1099						

RESULT 3
AAT99541
ID AAT99541 standard; cDNA; 2116 BP.
XX
AC AAT99541;
XX
DT 21-MAY-1998 (first entry)
XX
DE Human hyaluronan synthase cDNA clone 30C.
XX
KW Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
wound healing; vulnary; tissue repair; scar; keloid; therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 36..1772
FT /*tag= a
FT polyA_signal 2066..2071
FT /*tag= b
XX
PN WO9740174-A1.
XX
PD 30-OCT-1997.
XX
PF 17-APR-1997; 97WO-US006350.
XX
PR 22-APR-1996; 96US-00635552.
XX
PA (LEUK-) LEUKOSITE INC.
XX
PI Briskin MJ;
XX
DR WPI; 1997-549359/50.
DR P-PSDB; AAW26765.
XX
PT Human hyaluronan synthase - useful for recombinant production of
hyaluronic acid for wound healing, tissue repair and reducing
hypertrophic scar and keloid formation.
XX
PS Claim 3; Page 36-38; 58pp; English.
XX
CC cDNA clone 30C includes a coding region for hyaluronan synthase (HAS)
(see AAW26765), an enzyme involved in the synthesis of hyaluronan
(hyaluronic acid) and which has the ability to confer cell adhesion by
the lymphocyte receptor CD44. Clone 30C was isolated using an expression
cloning system developed to isolate cDNA clones that encode proteins that
confer adhesion of the murine T cell lymphoma TK1. A human mesenteric
lymph node expression library was constructed that, upon transfection
into CHO/P cells, yielded clone 30C that mediated rosetting of TK1 cells
to some of the transfectants. The isolated clone can be utilised in a
claimed method for producing HAS in host cells. Such host cells are used
in a claimed method for the production of hyaluronan. Hyaluronan is
useful for wound healing and tissue repair, and can reduce or prevent
hypertrophic scars and keloid formation. It is also used in eye surgery
as a replacement for vitreous fluid

Query Match	99.9%;	Score 1064;	DB 2;	Length 2116;
Best Local Similarity	100.0%;	Pred. No. 6.7e-180;		
Matches 1064;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAGACAGCAGGCGCGCCCAAGCCACTCTCTGAGCGCGCGCTGCTCCGGCTGGCC	60	
Db	36	ATGAGACAGCAGGCGCGCCCAAGCCACTCTCTGAGCGCGCGCTGCTCCGGCTGGCC	95	
QY	61	CGGAGGCTGTGACCATCGCTTGGCTTGTCTATCTTGGGCTCATGACCTGGGCTTAC	120	
Db	96	CGGAGGCTGTGACCATCGCTTGGCTTGTCTATCTTGGGCTCATGACCTGGGCTTAC	155	

```
QY 121 GCCGCGGGGTGCGCTGCTCCGATCGTACGGCTCTCGGCTTCGGCTCTACGGG 180
D 156 GCCGCGGGGTGCGCTGCTCCGATCGTACGGCTCTCGGCTTCGGCTCTACGGG 215
QY 181 GCTTCTCTTTTACGCGCACCTGCTGGCGCAGAGCTCTTTCGGCTACTGAGCACCGCGGG 240
D 216 GCTTCTCTTTTACGCGCACCTGCTGGCGCAGAGCTCTTTCGGCTACTGAGCACCGCGGG 275
QY 241 GTGGCGGGCGCGCGGGGGCGCTGATGAGCAACCGCGCGCAGTGTGGCGCTGACC 300
D 276 GTGGCGGGCGCGCGGGGGCGCTGATGAGCAACCGCGCGCAGTGTGGCGCTGACC 335
QY 301 ATCTCGGCTACAGAGGAGACCCCGGTACCTGCGCGCAGTGTGGCGCTGCGCGCGCC 360
D 336 ATCTCGGCTACAGAGGAGACCCCGGTACCTGCGCGCAGTGTGGCGCTGCGCGCGCC 395
QY 361 CTGCTGTATACCCCGCGCGCGCTGCGCGCTCTCATGCTGCTGATGAGCAACCGCGCGGAG 420
D 396 CTGCTGTATACCCCGCGCGCGCTGCGCGCTCTCATGCTGCTGATGAGCAACCGCGCGGAG 455
QY 421 GACCTCTACATGCTGACATGTTCCGCGAGGCTTTCGCTGACGAGGACCCCGCGCAGTAC 480
D 456 GACCTCTACATGCTGACATGTTTCGCGAGGCTTTCGCTGACGAGGACCCCGCGCAGTAC 515
QY 481 GTGTGGGACGCAACTACACAGCCCTGCGGAAACCGCGCGCGCGCGCGCTGCGCGCGCC 540
D 516 GTGTGGGACGCAACTACACAGCCCTGCGGAAACCGCGCGCGCGCGCGCTGCGCGCGCC 575
QY 541 CGAGGCTTATCGGAGGTGAGCGGAGGATCTCGGCGCGCTGGCAGTGGAGCGCTGGTG 600
D 576 CGAGGCTTATCGGAGGTGAGCGGAGGATCTCGGCGCGCTGGCAGTGGAGCGCTGGTG 635
QY 601 AGGACTCGAGGTGCTGTGCTGCGCAGCGCTGGCGCGCAACCGCGAGGTCAATGATAC 660
D 636 AGGACTCGAGGTGCTGTGCTGCGCAGCGCTGGCGCGCAACCGCGAGGTCAATGATAC 695
QY 661 ACAGCTTTCAGGCGCTCGGAGATTCGCTGAGTACGCTGAGTCTGAGTCTGGACACA 720
D 696 ACAGCTTTCAGGCGCTCGGAGATTCGCTGAGTACGCTGAGTCTGAGTCTGGACACA 755
QY 721 AGGTTGGACCCCATGCGCTGCTGAGCTCTGCTGCGGCTACTGAGAGGACCCCGGGTA 780
D 756 AGGTTGGACCCCATGCGCTGCTGAGCTCTGCTGCGGCTACTGAGAGGACCCCGGGTA 815
QY 781 GGGGCTGTGTGGGGAAGTGGGATCTTAACTCTGAGCTCTGGGCTGAGCTTCTTA 840
D 816 GGGGCTGTGTGGGGAAGTGGGATCTTAACTCTGAGCTCTGGGCTGAGCTTCTTA 875
QY 841 AGCAGCTCGGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCAC 900
D 876 AGCAGCTCGGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCAC 935
QY 901 TGTGTATCTGCTCAGCGGCTCTTAGGCTTATAGGATTAACCTCTTGCAGCAGTTT 960
D 936 TGTGTATCTGCTCAGCGGCTCTTAGGCTTATAGGATTAACCTCTTGCAGCAGTTT 995
QY 961 CTTGAGGCTGTGTAACAACAGAGTTCTTGGGTACCCACTGTGATTTTGGGATGACCGG 1020
D 996 CTTGAGGCTGTGTAACAACAGAGTTCTTGGGTACCCACTGTGATTTTGGGATGACCGG 1055
QY 1021 CACCTCACCAACCGCATCTCAGCATGGTGTATGCTACCAAGTA 1064
D 1056 CACCTCACCAACCGCATCTCAGCATGGTGTATGCTACCAAGTA 1099
```

RESULT 4

ADL13692

ID ADL13692 standard; DNA; 1737 BP.

XX AC ADL13692;

XX DT 06-MAY-2004 (first entry)

XX

```
DE Osteoarthritis-associated polymorphic nucleotide #224.
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX Homo sapiens.
XX WO2003054166-A2.
XX 03-JUL-2003.
XX 19-DEC-2002; 2002WO-US041225.
XX 20-DEC-2001; 2001US-0342603P.
XX (INCY-) INCYTE GENOMICS INC.
XX Jones KA, Schafer A;
XX WPI; 2003-559141/52.
XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polymorphic nucleotide encoding
XX a protein.
XX Disclosure; SEQ ID NO 224; 297pp; English.
XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polymorphic nucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polymorphic nucleotide encoding a protein listed in the specification. (Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences).
XX
XX Sequence 1737 BP; 232 A; 569 C; 604 G; 332 T; 0 U; 0 Other;
XX
Query Match 99.8%; Score 1062.4; DB 10; Length 1737;
Best Local Similarity 99.9%; Pred. No. 1.3e-179;
Matches 1063; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTCAGCCCGCCGCTGTCCGGCTGCGC 60
D 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTCAGCCCGCCGCTGTCCGGCTGCGC 60
QY 61 CGGAGGGTGTGACCATCGCTTCGCCCTGCTCATCTCGGCCCTCATGACCTGGGCTTAC 120
D 61 CGGAGGGTGTGACCATCGCTTCGCCCTGCTCATCTCGGCCCTCATGACCTGGGCTTAC 120
QY 121 GCCCGCGGGGTGCGCTGGCTCCGATCGCTACGGCCCTCTCGGCTTCGGGCTCTACGGG 180
D 121 GCCCGCGGGGTGCGCTGGCTCCGATCGCTACGGCCCTCTCGGCTTCGGGCTCTACGGG 180
QY 181 GCCTTCCTTTACGCGCACCTGGTGGCGCAGAGCTCTTTCGGGTACTGAGACACCGGGCG 240
D 181 GCCTTCCTTTACGCGCACCTGGTGGCGCAGAGCTCTTTCGGGTACTGAGACACCGGGCG 240
QY 241 GTGGCGGGCGCGCGGGGGCGCTGATGAGCAACCGCGCGCAGTGTGGCGCTGACC 300
D 241 GTGGCGGGCGCGCGGGGGCGCTGATGAGCAACCGCGCGCAGTGTGGCGCTGACC 300
QY 301 ATCTCGGCTACAGAGGAGACCCCGGTACCTGCGCGCAGTGTGGCGCTGCGCGCGCC 360
```


Db 301 ATCTCCGCTTACAGAGGACCCCGCTACCTGCGCCAGTGCCTGGCGCTCCGCGCGCC 360
Qy 361 CTGCTGTATACCGCGCGCGCTGGCGTCTCTCATGTGTGGTGGCAACCGCGCCGAG 420
Db 361 CTGCTGTATACCGCGCGCGCTGGCGTCTCTCATGTGTGGTGGCAACCGCGCCGAG 420
Qy 421 GACCTCTACATGGTGCACATGTTCCGAGAGTCTTCGCTGACGAGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGGTGCACATGTTCCGAGAGTCTTCGCTGACGAGACCCCGCCACGTAC 480
Qy 481 GTGTGGGACGGCACTACACAGCCCTGGGACCCCGCGCGCGCGCGCGCGCGCGCC 540
Db 481 GTGTGGGACGGCACTACACAGCCCTGGGACCCCGCGCGCGCGCGCGCGCGCGCC 540
Qy 541 GGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAGCGCTGGTG 600
Db 541 GGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAGCGCTGGTG 600
Qy 601 AGGACTCGAGGTGGTGTGGTGGCGAGCGCTGGCGCGCAAGCGGAGGTCTATGTAC 660
Db 601 AGGACTCGAGGTGGTGTGGTGGCGAGCGCTGGCGCGCAAGCGGAGGTCTATGTAC 660
Qy 661 ACAGCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGACTCGACACA 720
Db 661 ACAGCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGACTCGACACA 720
Qy 721 AGTTGGACCCCATGGCACTGCTGAGCTCTGGTGGTACTGGAGGAGCCCGCGGTA 780
Db 721 AGTTGGACCCCATGGCACTGCTGAGCTCTGGTGGTACTGGAGGAGCCCGCGGTA 780
Qy 781 GGGGCTGTGTGGGAGCGTGGGATCTTAACCTCTGGAATCTCTGGGTGAGTTCCTA 840
Db 781 GGGGCTGTGTGGGAGCGTGGGATCTTAACCTCTGGAATCTCTGGGTGAGTTCCTA 840
Qy 841 AGCAGCTGCATCTAGTGGTACCTTCAATGTGGAGCGGCTTGTGAGACTACTTCCAC 900
Db 841 AGCAGCTGCATCTAGTGGTACCTTCAATGTGGAGCGGCTTGTGAGACTACTTCCAC 900
Qy 901 TGTGTATCTGATCAGCGGCTCTAGGCTTATAGGAATAAATCTTTCAGAGATT 960
Db 901 TGTGTATCTGATCAGCGGCTCTAGGCTTATAGGAATAAATCTTTCAGAGATT 960
Qy 961 CTGAGGCTGGTACAAACAGAGTTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
Db 961 CTGAGGCTGGTACAAACAGAGTTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
Qy 1021 CACCTCACCAACCGCATCTAGCATGGGTTATGCTACCAAGTA 1064
Db 1021 CACCTCACCAACCGCATCTAGCATGGGTTATGCTACCAAGTA 1064

RESULT 5
ADL13694

ID ADL13694 standard; DNA; 2087 BP.

XX AC ADL13694;
XX AC

XX DT 06-MAY-2004 (first entry)

XX DE Osteoarthritis-associated polymorphic nucleotide #226.

XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.

XX OS Homo sapiens.

XX PN WO2003054166-A2.

XX PD 03-JUL-2003.

XX PF 19-DEC-2002; 2002WO-US041225.

XX XX

PR 20-DEC-2001; 2001US-0342603P.

XX (INCY-) INCYTE GENOMICS INC.
XX

PI Jones KA, Schafer A;

XX WPI; 2003-559141/52.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

20-DEC-2001; 2001US-0342603P.
(INCY-) INCYTE GENOMICS INC.
Jones KA, Schafer A;
WPI; 2003-559141/52.
Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.
Disclosure; SEQ ID NO 226; 297pp; English.
The invention relates to a method of determining susceptibility of an
individual to joint space narrowing and/or osteophyte development and/or
joint pain comprising identifying whether the individual has at least one
polymorphism in a polynucleotide encoding at least one of the protein
listed in the specification. The methods, composition and agent are
useful for modulating the susceptibility of an individual to joint space
narrowing and/or osteophyte development and/or joint pain that is
associated with a disease, preferably osteoarthritis. The cell line and
the non-human animal are useful for screening for an agent for diagnosing
an individual having susceptibility to joint space narrowing and/or
osteophyte development and/or joint pain. This sequence corresponds to
the polynucleotide encoding a protein listed in the specification. (Note:
The sequence data for this patent did not form part of the printed
specification but was obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences).

Sequence 2087 BP; 302 A; 650 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 99.8%; Score 1062.4; DB 10; Length 2087;

Best Local Similarity 99.9%; Pred. No. 1.3e-179;

Matches 1063; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGACAGCAGGACGGCCCAAGCCCACTCTCTGACGACCCCGCTGTCTGGGCTTGCC 60
Db 36 ATGAGACAGCAGGACGGCCCAAGCCCACTCTCTGACGACCCCGCTGTCTGGGCTTGCC 95
Qy 61 CGAGGGTGTGACATCGCCTTCCGCTGTCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 96 CGAGGGTGTGACATCGCCTTCCGCTGTCTCATCTCTGGGCTCATGACCTGGGCTTAC 155
Qy 121 GCCCGCGGGTGGCGCTGGCCTCCGATCGCTACGCGCTCTGGCTTCCGGCTTCTACGG 180
Db 156 GCCCGCGGGTGGCGCTGGCCTCCGATCGCTACGCGCTCTGGCTTCCGGCTTCTACGG 215
Qy 181 GCCTTCTCTTTCAGGCGCACCTGGTGGCGCAGAGCCTCTTCGCGTACCTGGAGCACCGCGG 240
Db 216 GCCTTCTCTTTCAGGCGCACCTGGTGGCGCAGAGCCTCTTCGCGTACCTGGAGCACCGCGG 275
Qy 241 GTGGCGCGCGCGCGCGCGCGCGCGCTGGATGACAGCACCGCGCGAGTGTGGCGCTGACC 300
Db 276 GTGGCGCGCGCGCGCGCGCGCGCGCTGGATGACAGCACCGCGCGAGTGTGGCGCTGACC 335
Qy 301 ATCTCCGCTTACAGGAGGACCCCGCTTACCTGGCGCAGTGCCTGGCGCTCCGCGCGCC 360
Db 336 ATCTCCGCTTACAGGAGGACCCCGCTTACCTGGCGCAGTGCCTGGCGCTCCGCGCGCC 395
Qy 361 CTGCTGTATACCGCGCGCGCGCTGGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420
Db 396 CTGCTGTATACCGCGCGCGCGCTGGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 455
Qy 421 GACCTCTACATGGTGCACATGTTCCGAGAGTCTTCGCTGACAGAGACCCCGCCACGTAC 480
Db 456 GACCTCTACATGGTGCACATGTTCCGAGAGTCTTCGCTGACAGAGACCCCGCCACGTAC 515
Qy 481 GTGTGGGACGGCAACTACCAAGCCCTTGGGAAACCCCGGGCGGGCGCGGTGGGCGCC 540
Db 516 GTGTGGGACGGCAACTACCAAGCCCTTGGGAAACCCCGGGCGGGCGCGGTGGGCGCC 575

QY 541 GGAGCTATCGGAGGTGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAGCGCTGGTG 600
 DB |||||
 QY 576 GGAGCTATCGGAGGTGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAGCGCTGGTG 635
 DB |||||
 QY 601 AGGACTCGAGGTGCGTGTGCGCGCAGCGCTGGCGCGCAAGCGCGAGGTCAATGAC 660
 DB |||||
 QY 636 AGGACTCGAGGTGCGTGTGCGCGCAGCGCTGGCGCGCAAGCGCGAGGTCAATGAC 695
 DB |||||
 QY 661 ACAGCCTTCAGGCGCTGGAGATTCGGTGGACTACGTGCGAGGTGTGACTCGGACACA 720
 DB |||||
 QY 696 ACAGCCTTCAGGCGCTGGAGATTCGGTGGACTACGTGCGAGGTGTGACTCGGACACA 755
 DB |||||
 QY 721 AGTTGGACCCCATCGCACTGCTGAGCTCGTGGGCTACTGGACGAGGACCCCGGGTA 780
 DB |||||
 QY 756 AGTTGGACCCCATCGCACTGCTGAGCTCGTGGGCTACTGGACGAGGACCCCGGGTA 815
 DB |||||
 QY 781 GGGGCTGTGTGGGAGCTGCGGATCCTTAACCTCTGGACTCTCTGGGTGAGTTCCTA 840
 DB |||||
 QY 816 GGGGCTGTGTGGGAGCTGCGGATCCTTAACCTCTGGACTCTCTGGGTGAGTTCCTA 875
 DB |||||
 QY 841 AGCAGCCTCGCATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGACTACTTCCAC 900
 DB |||||
 QY 876 AGCAGCCTCGCATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGACTACTTCCAC 935
 DB |||||
 QY 901 TGTGTATCTGCATCAGCGGCTCTAGGCTTATAGGAATAACCTCTTGCAGCAGTTT 960
 DB |||||
 QY 936 TGTGTATCTGCATCAGCGGCTCTAGGCTTATAGGAATAACCTCTTGCAGCAGTTT 995
 DB |||||
 QY 961 CTTGAGGCTGGTACAAACAGAGTTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
 DB |||||
 QY 996 CTTGAGGCTGGTACAAACAGAGTTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1055
 DB |||||
 QY 1021 CACCTCACCAACCGCATCTGAGCATGGGTATGCTACCAAGTA 1064
 DB |||||
 QY 1056 CACCTCACCAACCGCATCTGAGCATGGGTATGCTACCAAGTA 1099
 DB |||||

RESULT 6

AAAT96713
 ID AAAT96713 standard; DNA; 2117 BP.
 AC AAAT96713;
 XX
 DT 22-APR-1998 (first entry)
 XX
 DE Human hyaluronate synthetase coding sequence.
 XX
 KW Hyaluronate synthetase; human; hyaluronic acid; drug preparation;
 XX cosmetic preparation; gene therapy; carcinogenesis; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 149..1780
 FT /*tag= a
 XX
 FN W09738113-Al.
 XX
 PD 16-OCT-1997.
 XX
 PF 31-MAR-1997; 97WO-JP001111.
 XX
 PR 05-APR-1996; 96JP-00084326.
 PR 30-APR-1996; 96JP-00109663.
 XX
 PA (SEK) SEIKAGAKU CORP.
 XX
 PI Itano N, Kimata K;
 XX
 DR WPI; 1997-512726/47.
 DR P-PSDB; AAW36503.
 XX
 PT DNA encoding human hyaluronate synthetase - for industrial scale

PT production of hyaluronic acid used in generating anti-carcinogenic drugs
 PT or for cosmetics.
 XX
 PS Claim 3; Page 23-27; 35pp; Japanese.
 XX
 CC This sequence encodes a human hyaluronate synthetase, and is the coding
 CC sequence of the invention. The encoded enzyme is useful for industrial
 CC scale production of hyaluronic acid for use in the preparation of drugs
 CC and cosmetics. The drugs can also be used in compositions for the
 CC treatment of disorders involving the lowering of hyaluronic acid
 CC production. The peptides may be used for the preparation of antibodies
 CC recognising hyaluronate synthetase, e.g. for diagnostic purposes.
 CC Antisense DNA or RNA corresponding to the DNA can be used in gene therapy
 CC treatment of carcinogenesis
 XX
 SQ Sequence 2117 BP; 306 A; 673 C; 703 G; 435 T; 0 U; 0 Other;
 Query Match 97.8%; Score 1042; DB 2; Length 2117;
 Best Local Similarity 99.1%; Pred. No. 5.4e-176;
 Matches 1048; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 7 CAGCAGGACGCGCCAAAGCCACCTCTGCAGCCCGCGCTGCTCGGCTCGGCCGAGG 66
 DB |||||
 QY 50 CAGCAGGACGCGCCAAAGCCACCTCTGCAGCCCGCGCTGCTCGGCTCGGCCGAGG 109
 DB |||||
 QY 67 GTGCTGACCATCGCCTTTCGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTACGCGCC 126
 DB |||||
 QY 110 GTGCTGACCATCGCCTTTCGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTACGCGCC 169
 DB |||||
 QY 127 GGGGTGCGCTGGCTTCGATCGCTACGGCTCTCGGCTTTCGGCTTACGGGCTTTC 186
 DB |||||
 QY 170 GGGGTGCGCTGGCTTCGATCGCTACGGCTCTCGGCTTTCGGCTTACGGGCTTTC 229
 DB |||||
 QY 187 CTTTCAGCGACCTGTGTGGCGCAGAGCTCTTTCGGTACCTGGAGACACCGCGGGTGGCG 246
 DB |||||
 QY 230 CTTTCAGCGACCTGTGTGGCGCAGAGCTCTTTCGGTACCTGGAGACACCGCGGGTGGCG 289
 DB |||||
 QY 247 GCGCGCGCGGGCGCTGGATGACGACACCGCGCGCAGTGTGGCTTACCATCTCC 306
 DB |||||
 QY 290 GCGCGCGCGGGCGCTGGATGACGACACCGCGCGCAGTGTGGCTTACCATCTCC 349
 DB |||||
 QY 307 GCCTACGAGGAGACCCCGCTTACCTGCGCAGTGTGGCTTTCGGCTTTCGGCTTTC 366
 DB |||||
 QY 350 GCCTACGAGGAGACCCCGCTTACCTGCGCAGTGTGGCTTTCGGCTTTCGGCTTTC 409
 DB |||||
 QY 367 TACCGCGCGCGCTGCGGCTCTCATGTGTGTGATGACCAACCGCGCGGAGACCTC 426
 DB |||||
 QY 410 TACCGCGCGCGCTGCGGCTCTCATGTGTGTGATGACCAACCGCGCGGAGACCTC 469
 DB |||||
 QY 427 TACATGTGTGACATGTTCCGCGAGGTCTTCTGCTACGAGGACCCCGCAGCTGTGTGG 486
 DB |||||
 QY 470 TACATGTGTGACATGTTCCGCGAGGTCTTCTGCTACGAGGACCCCGCAGCTGTGTGG 529
 DB |||||
 QY 487 GACGCGCAACTACCAACGAGCTTGGGAAACCGCGCGCGGCGGCTGGGCGCGGAGCC 546
 DB |||||
 QY 530 GACGCGCAACTACCAACGAGCTTGGGAAACCGCGCGGCGGCGGCTGGGCGCGGAGCC 589
 DB |||||
 QY 547 TATCGGAGGTGTGAGGCGGAGATTCCTGCGGCGCTGGCAGTGTGGAGGCGCTGGTGGAGCT 606
 DB |||||
 QY 590 TATCGGAGGTGTGAGGCGGAGATTCCTGCGGCGCTGGCAGTGTGGAGGCGCTGGTGGAGCT 649
 DB |||||
 QY 607 CGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTTCATGTACACAGCC 666
 DB |||||
 QY 650 CGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTTCATGTACACAGCC 709
 DB |||||
 QY 667 TTCAAGCGCGCTCGGAGATTCTGGTGGACTACGTGAGGTCTGTGTGCTCGGACACAAGGTTG 726
 DB |||||
 QY 710 TTCAAGCGCGCTCGGAGATTCTGGTGGACTACGTGAGGTCTGTGTGCTCGGACACAAGGTTG 769
 DB |||||
 QY 727 GACCCCATGCGACTGTGAGCTGCTGGGTTACTGACGAGGACCCCGGTTAGGGCT 786
 DB |||||
 QY 770 GACCCCATGCGACTGTGAGCTGCTGGGTTACTGACGAGGACCCCGGTTAGGGCT 829
 DB |||||

QY 787 GTTGTGGGACGTGGGATCTTAACCTCTGGACTCTGGTCTGCTTAAGCAGC 846
 Db 830 GTTGTGGGATGTCCGGATCTTAACCTCTCGACTCTGGTCTGCTTAAGCAGC 889
 QY 847 CTGCGATCTGGGTAGCCTTCAATGTGGAGCGGCTTGTACAGCTACTTCCACTGTGA 906
 Db 890 CTGCGATCTGGGTAGCCTTCAATGTGGAGCGGCTTGTACAGCTACTTCCACTGTGA 949
 QY 907 TCTTCATCAGCGGCTCTTAGCCCTATATAGGAATAACCTCTTGACGAGTTCCTTGAG 966
 Db 950 TCTTCATCAGCGGCTCTTAGCCCTATATAGGAATAACCTCTTGACGAGTTCCTTGAG 1009
 QY 967 GCCTGTACACAGAGTTCCTGGTACCCACTGTACTTTTGGGATGACCGGCACTTC 1026
 Db 1010 GCCTGTACACAGAGTTCCTGGTACCCACTGTACTTTTGGGATGACCGGCACTTC 1069
 QY 1027 ACCAACCGCATGCTCAGCATGGTGTATGCTACCAAGTA 1064
 Db 1070 ACCAACCGCATGCTCAGCATGGTGTATGCTACCAAGTA 1107

RESULT 7

ID ABZ76734 standard; cDNA; 1752 BP.
 XX

AC ABZ76734;
 DT

XX 01-APR-2000 (first entry)
 XX

DE Mouse hyaluronan synthase 1 encoding cDNA SEQ ID NO:1.
 XX

KW Mouse; hyaluronan synthase; HAS; HAS1; enzyme; ophthalmological;
 KW osteopathic; antiarthritic; gene therapy; angiogenesis inhibitor;
 KW hyaluronic acid; dihydrazide; bioconjugate; dry eye syndrome; eye;
 KW osteoarthritis; gene; ss.
 XX

OS Mus sp.
 XX

FH Key Location/Qualifiers
 FT 1..1752
 FT CDS /tag= a
 FT /product= "hyaluronan synthase 1 (HAS1)"
 FT

PN WO2003006068-A1.
 XX

XX 23-JAN-2003.
 PD

XX 10-JUL-2001; 2001WO-US021785.
 PF

XX 10-JUL-2001; 2001WO-US021785.
 PR

XX (CLEA-) CLEAR SOLUTIONS BIOTECH INC.
 XX

XX Dehazya P, Chen W;
 XX

XX WPI; 2003-221664/21.
 DR

XX P-PSDB; ABP96028.
 DR

PT Novel dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
 PT for treating dry eye syndrome, has derivatized hyaluronic acid
 PT crosslinked to nucleic acid encoding protein with hyaluronan synthase
 PT activity.
 PT

XX Claim 19; Page 56; 62pp; English.
 PS

XX The present invention describes a dihydrazide derivatised hyaluronic acid
 CC (HA)/nucleic acid (NA) bioconjugate (I), with derivatised HA linked to NA
 CC encoding protein with hyaluronan synthase (HAS) activity, where NA has
 CC sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659
 CC or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding
 CC proteins with 66-90% SI to a sequence of 583, 552 or 554 amino acids (see
 CC ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and
 CC antiarthritic activities, and can be used in gene therapy and as an

CC inhibitor of angiogenesis, and as an inducer of expression of (HA) in
 CC human corneal epithelial cell. (I) is useful for transfecting a cell of
 CC an eye with nucleic acid, where the nucleic acid comprises (ABZ76734 to
 CC ABZ76736), by contacting the cell with (I) comprising the nucleic acid.
 CC (I) is useful for treating dry eye syndrome in an individual. (I) is
 CC useful in gene therapy applications for the treatment of a variety of
 CC medical conditions including dry eye syndrome or other medical conditions
 CC where an increase in the production of (HA) in the eye would be
 CC therapeutic (e.g., osteoarthritis of the articular joints). (I) is also
 CC useful for inhibiting angiogenesis for the treatment of macular
 CC degeneration or genes related to lipid biosynthesis that helps to restore
 CC the lipid component of the tear film, and as reagents for in vitro
 CC transformation of any cell, preferably a eukaryotic cell, more preferably
 CC a human eye cell. The present sequence encodes mouse HAS1 which is used
 CC in the exemplification of the present invention
 XX

SQ Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;

Query Match 73.9%; Score 786.6; DB 8; Length 1752;
 Best Local Similarity 84.9%; Pred. No. 1.2e-130;
 Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;

QY 10 CAGGACGGCCCAAGCCCACTCTCGACGCGCGCTGCTCGGCGCTGGCGGAGGTG 69
 Db 7 CAGGACATGCCAAAGCCCTCAGAGCGCGCTTGTCTCTGCGCTGGCGGAGCA 66
 QY 70 CTGACCATCGCTTCGCCCTGCTCATCTGGGCTCATGACTGGGCTAGCGCCCGGG 129
 Db 67 CTCAGATCATCTTGGCTGCTCATCTGGGCTCATGACTGGGCTAGCGCCGAGGC 126
 QY 130 GTGCGCTGGCTCCGATCGCTACGGCTCTCTGGCTTCCTGGGCTCTACGGGCTTCCTT 189
 Db 127 GTTCTCTGGCTTCAGATCGCTATGAGCTCTCTGGCTTCCTGGGCTCTATGGGCTTCCTC 186
 QY 190 TCAGGCGACCTGTGGGCGCAGAGCTCTTGGGTACCTTGGAGCACCGCGGGTGGCGG 249
 Db 187 AGCGCACACCTAGTGGCACAGAGCTCTTCTGCTTACCTGGAGCACCGGAGGGTGGCAGCG 246
 QY 250 GCGGCGCG-----GGGCGCGTGGATGAGCAGCCAGCGCGCAGTGTGGCG 294
 Db 247 GCTGGCGGCGCTCTTGGCGAAGGGGCTCTGGATGGGCTCTGACGAGCGTGGCA 306
 QY 295 CTGACCATCTCCGCTTACGAGAGGACCCCGCTACCTGCGCCAGTGTGCGCTCGCGC 354
 Db 307 CTCACCATCTCAGCTTACCAAGAGGATCCCGCTTACCTGCGCCAGTGTGACCTCCGCG 366
 QY 355 CGCGCTCTGTGTACCGCGCGCGGCTGCGGCTCTCATGTGGTGGATGGCAACCGC 414
 Db 367 CGCGCTTGTGTATCCCGCACACAGAGGTACGCGTGTCTCATGTGGTGGAGCGCAACCGC 426
 QY 415 GCCGAGGACCTCTACATGTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGC 474
 Db 427 GCTGAGGATCTGTATGTGTGACATGTTCCGAGAGTCTTCGCGCATGAGAGACCCCGC 486
 QY 475 ACGTACGTGTGGGACGGCAACTACACAGCCCTGGGAACCCCGCGCGG-----CGGCGCG 531
 Db 487 ACTTATGTGGGATGGCAACTACCATCAGCCCTGGGAACCCAGCGAGGCTACGGGCGCT 546
 QY 532 GTGGCGCGGAGGCTATCGGGAGGTGGAGCGGAGGATCTCTGGCGGCTGCGAGTGGAG 591
 Db 547 GTCGCTGAAGTGTGCTTACCGGAGGTGGAGCGGAGGACCCCGCGCGGTGTGGCGTGGAG 606
 QY 592 GCGCTGTGTGAGGACTCGCAGGTGTGTGCTGCGCAGCGCTGGGGCGGCGAAGCGCAG 651
 Db 607 GCGCTGTGTGAGAACACGCGAGGTGTGTGCTGCGCTGAGCGTGTGGGCGGCAACGCTGAG 666
 QY 652 GTCATGTACACAGGCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAGTCTGTGAC 711
 Db 667 GTCATGTACACAGCTTCAAGGCACTGGCGGACTCCGTTGGGACTACGTGCAGTCTGTGAC 726
 QY 712 TCGGACACAAAGTTGGACCCCATGTGCTGAGCTGTGTGCGGGTACTGTGACGAGGAC 771
 Db 727 TCAGACACAAAGCTAGACCCCATGTGCTGAGCTGTGTGCGAGTGTGTGATGAAGAC 786

QY 772 CCCGGTAGGGCTGTTGGTGGGACCTGCGGATCCCTTACCCCTCGACTCCTGGTCTC 831
Db 787 CCCGGTAGGGCTGTTGGAGGGATGTGAGGATCCCTTAAACCTCTGGACTCCTGGGTCTC 846
QY 832 AGCTTCTTAAGCAGCTGCTGATCTAGGCTTCAATGTGGAGCGGCTTGTGAGAGC 891
Db 847 AGCTTCTTGAGCAGCTTTCGATCTAGGCTTCAATGTGGAGCGGCTTGTGAGAGC 906
QY 892 TACTTCCACTGTGATCTGCAATCAGCGGTCTCTAGGCTTATATAGGAATAACCTCTTG 951
Db 907 TACTTCCACTGTGCTGCAATCAGTGTGCTCTGGGTCTATACAGAAAACAATCTCCTG 966
QY 952 CAGCAGTTTCTTGAGGCTGTGACACAGAGTTTCTGGGTACCCACTGTACTTTTGGG 1011
Db 967 CAGCAGTTTCTTGAGGCTGTGACACAGAGTTTCTGGGTACCCACTGTACATTTGGG 1026
QY 1012 GATGACCGGCACCTCACCAACCGCATGCTCAGCATGGGTATGTACCAAGTA 1064
Db 1027 GATGACCGGCACCTCACCAACCGCATGCTTAGCATGGGTATGTACCAAGTA 1079

RESULT 8

AAD59442
ID AAD59442 standard; DNA; 1752 BP.
XX AAD59442;

DT 18-DEC-2003 (first entry)

DE Mouse hyaluronan synthase (HAS) 1 DNA.

XX Mouse; hyaluronic acid; HA; hyaluronan synthase; HAS; dry eye syndrome;
KW eye cell; osteoarthritis; gene therapy; gene; ds.

XX Mus sp.

FH Key Location/Qualifiers

FT CDS 1..1752

FT /*tag= a

FT /product= "Mouse hyaluronan synthase 1"

XX US2003087850-A1.

XX 08-MAY-2003.

XX 10-JUL-2001; 2001US-00902939.

XX 10-JUL-2001; 2001US-00902939.

XX (DEHA/) DEHAZYA P.

XX (CHEN/) CHEN W.

XX Dehazy P, Chen W;

XX WPI; 2003-755151/71.

XX P-P5DB; AAE39152.

XX Dihydrate derivatized hyaluronic acid/nucleic acid bioconjugate
PT comprising derivatized hyaluronic acid cross linked to nucleic acid
PT encoding hyaluronan synthase useful for treating dry eye syndrome.

XX Claim 1; Page 12-13; 31pp; English.

XX The invention relates to dihydrate derivatized hyaluronic acid (HA)/
CC nucleic acid bioconjugate comprising derivatized HA cross linked to
CC nucleic acid encoding hyaluronan synthase (HAS). The invention is useful
CC for transfecting an eye cell of an individual. It is useful for treating
CC dry eye syndrome and osteoarthritis of the particular joints. The
CC invention is also useful in gene therapy. The present sequence is mouse
XX HAS1 DNA

XX Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;

Query Match 73 9%; Score 786.6; DB 10; Length 1752;
Best Local Similarity 84.9%; Pred. No. 1.2e-130;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;
QY 10 CAGGACGCGCCCAAGACCCACCTCTCTGACAGCCCGCGCTGCTCCGGCTTGGCCGAGGGTGTG 69
Db 7 CAGGACATGCCAAGACCCCTCAGAGGAGCGGGTGTGCTCTGCGCTTGGCCGAGGGAGCA 66
QY 70 CTGACCATCGCTTTCGCGCTCTCATCTGCGGCTCTATGACCTTGGGCTCTACGCGCCCGGG 129
Db 67 CTCAGATCATCTTTTGGCCCTCATCTGCGGCTCTATGACCTTGGGCTCTACGCGCAGGC 126
QY 130 GTGCGCTGCGCTCCGATCGCTAGCGGCTCTGCGCTTTCGGGCTCTACGCGGCTTCTCTT 189
Db 127 GTTCTCTTGGCTTTCAGATCGCTATGAGACTCTCTGCGCTTTCGGGCTCTATGCGGCTTCTC 186
QY 190 TCAGCGCACCTGTGCGGCGCAGAGCTCTTTCGCTACCTTGGAGCACCGCGGCTTGGCGCG 249
Db 187 AGCGCACACCTAGTGGCACAGAGCTCTTTCGCTTACCTTGGAGCACCGGAGGTTGGCAGG 246
QY 250 GCGGCGCG-----GGGCGCTGTGATGACGCCACCGCGCGCAGTGTGGCG 294
Db 247 GCTGCGGCGCTCTTTCGCGAAGGGGCGCTTGGATGCGGCCACTGTGACGAGCGTGGCA 306
QY 295 CTGACCATCTCCGCTTACCGAGGAGACCCCGCTACCTGCGCCAGTGTCTGCGCTCGGCC 354
Db 307 CTCACCATCTCAGCTTACCGCACACGAGGTTCACGCTGCTCATGTGTGGAGCGGCAACGCG 366
QY 355 CGCGCTCTGTGTACCGCGCGCGCTGCGCTCTCTCATGTGTGGATGGCAACGCG 414
Db 367 CGCGCTTGTGTACCGCACACGAGGTTCACGCTGCTCATGTGTGGAGCGGCAACGCG 426
QY 415 GCGGAGACCTCTACATGTGTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCC 474
Db 427 GCTGAGGATCTGTATGATGTGTGACATGTTCCGAGAGTCTTTCGCGCATGAGGACCCCGCC 486
QY 475 AGCTACGTGTGGAGCGGCACTTACACAGCCCTTGGGAAACCCCGCGCGG---CGGGCGCG 531
Db 487 ACTTATGTGTGGATGGCACTTACATCAGCCCTTGGGAAACCCAGCGGCTTACGCGGCT 546
QY 532 GTGGCGCGCGGAGCTTATCGGGAGGTGGAGCGGAGGATCTCTGGCGGCTTGGCAGTGGAG 591
Db 547 GTCGTGAAGGTGCTTACCGGAGGTGGAGCGGAGGACCCCGCGGCTTGGCGTGGAG 606
QY 592 GCGCTGTGAGGACTCGCAGGTGTGTGCTGCGCGCAGCGCTGGGGCGGCAAGCGCAG 651
Db 607 GCGCTGTGAGAACACACGAGGTGTGTGCTGCGCTCAGCGCTTGGGGCGGCAACGCTGAG 666
QY 652 GTCATGTACACAGCTTTCAGGCGCTCGGAGATTCGGTGGACTACGTGAGCTCTGTGAC 711
Db 667 GTCATGTACACAGCTTTCAGGCGCTCGGCGACTCCGTTGGACTACGTGAGCTCTGTGAC 726
QY 712 TCGGACACAGGTTGGACCCCATGCGACTGTGGAGCTCTGTGCGGCTTCTGGACGAGGAC 771
Db 727 TCAGACACAGACTAGACCCCATGCGACTGTGGAGCTTGTGCGAGTGTGGATGAAGAC 786
QY 772 CCGCGGTAGGGCTGTGTGGGAGCGTGGGATTCCTTAAACCTCTGGACTCTCGGTC 831
Db 787 CCGCGGTAGGGCTGTGTGGAGGGGATGTGAGGATTCCTTAAACCTCTGGACTCTCGGTC 846
QY 832 AGCTTCTTAGCAGCTTTCGATCTGCGTACCTTCAATGTGGAGCGGCTTGTTCAGAGC 891
Db 847 AGCTTCTTAGCAGCTTTCGATCTGCGTACCTTCAATGTGGAGCGGCTTGTTCAGAGC 906
QY 892 TACTTCCACTGTGTATCTCTGATCAGCGGCTCTCTAGGCTTATATAGGAATAACCTCTTG 951
Db 907 TACTTCCACTGTGTCTCTGATCAGTGTCTCTGGGTCTTATACAGAAACAATCTCCTG 966
QY 952 CAGCAGTTTCTTGAGGCTGTGTACCAACAGAGTTTCTGGGTACCCACTGTACTTTTGGG 1011
Db 967 CAGCAGTTTCTTGAGGCTGTGTACCAACAGAGTTTCTGGGTACCCACTGTACATTTGGG 1026

QY 1012 GATGACGGGACCTCACCACCGCATGCTCAGCATGGGTTATGCTACCAAGTA 1064
 Db 1027 GATGACGGGACCTCACCACCGCATGCTTAGCATGGGCTATGCTACCAAGTA 1079

RESULT 9

AAT91655
 ID AAT91655 standard; DNA; 2102 BP.

AC AAT91655;
 DT 17-OCT-2003 (revised)
 DT 19-DEC-1997 (first entry)
 XX Mouse hyaluronate synthase genomic DNA.
 XX Hyaluronate synthase; hyaluronic acid; polysaccharide; ds.
 XX Mus sp; (strain FM3A).

XX Key Location/Qualifiers
 FT 49..1800
 FT /tag= a
 FT /product= "Hyaluronate_synthase"

XX JP09224674-A.

XX 02-SEP-1997.

XX 26-FEB-1996; 96JP-00038336.

XX 26-FEB-1996; 96JP-00038336.

XX (KAGG) KAGAKU GIJUTSUCHO CHOKAN KANBO.

XX WPI: 1997-484102/45.

XX P-PSDB; AAW30704.

XX Hyaluronate synthase isolated from mouse cells - useful for large-scale production of hyaluronic acid.

XX Claim 2; Page 10-13; 15pp; Japanese.

CC A mouse-derived cell strain, FM3A P15-A, was cultured in Eagle minimal medium containing 10 % heat-inactivated bovine serum, twice concentration of amino acids and vitamins and penicillin and streptomycin at 37 degrees Celsius. The culture was subjected to immobilised erythrocyte exclusion to examine the extent of extracellular formation of hyaluronic acid matrix. Cells which showed high formation were recovered and named FM3A HAI. Cells synthesising no hyaluronic acid were recovered and named HAS-. The HAS- cell in which polyoma large T antigen was expressed was prepared (HAS- P cell). Poly(A)+ RNA was isolated from FM3A HAI and cDNA was prepared and was used for constructing a library in HAS- P cells. Cells having hyaluronic acid synthetic activity were selected from the transformants and plasmid DNA was recovered and amplified in E.coli. The resulting genomic DNA sequence codes for hyaluronate synthase having a sequence of 583 amino acids which is used for large-scale production of hyaluronic acid, a component of drugs and cosmetics. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;

Query Match 73.9%; Score 786.6; DB 2; Length 2102;
 Best Local Similarity 84.9%; Pred. No. 1.2e-130;
 Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;

QY 10 CAGGACGGCCCAAGCCACTCTCGAGCCCGCGCTGCTCGGCTGCGCCGGAGGGTG 69
 Db 55 CAGGACATGTCACCAAGCCCTCAGAGGCGCGTGTCTGCTGCGCTGGCCAGGCGAGCA 114
 QY 70 CTGACCATCGGCTTGGCCGTGCTATCTCTGGGCTCATGACCTGGGCTAGCCCGCGGG 129
 Db 115 CTCAGATCATCTTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTAGCCCGCAGGC 174

QY 130 GTGCGCTGGCTCCGATCGCTACGGCCCTCTGGCCCTTCCGGCTCTACGGGGCCCTTCCTT 189
 Db 175 GTTCTCTGGCTTCAGATCGCTATGGACTCTCTGGCTTTGGCTCTATGGGGCATTCCTC 234
 QY 190 TCAGGCACTGTGGTGGCGCAGAGCCTCTTTCGGTACCTGGAGCACCGCGGGTGGCGGG 249
 Db 235 AGCGCACACCTAGTGGCACAGAGCCTCTTTCGCTTACCTGGAGCACCGAAGGGTGGCAGCG .294
 QY 250 GCGGCGCG-----GGGCGCGCTGGATGAGCCGACCCAGCCGCGCAGTGTGGCG 294
 Db 295 GCTGCGCGGCGCTCTTGGCGAAGGGGCCCTTGGATGCGGCCACTGCGACGAGCGTGGCA 354
 QY 295 CTGACCATCTCGGCTTACAGAGGACCCCGCGTACCTGCGCCAGTGCCTGCGCTCGCGC 354
 Db 355 CTCACCATCTCAGCCTACCAAGAGGATCCCGCTTACCTGCGCCAGTGTCTGACCTCGCG 414
 QY 355 CGCGCCCTGTGTATCCCGCGCGCGGCTGCGGCTCTCATGTGGTGGTGGTGGTGGTGGTGG 414
 Db 415 CGCGCTTGTGTATCCCGCACACGAGGTTACGCGTGTCTCATGTGGTGGTGGTGGTGGTGGTGG 474
 QY 415 GCGGAGGACCTCTACATGTCGACATGTTCCGCGAGGTTCTTCGCTGACGAGGACCCCGCC 474
 Db 475 GCTGAGGATCTGTATCATGTTGAGCATGTTCCGAGAAGTCTTTCGCGATGAGGACCCCGCC 534
 QY 475 ACGTACGTGTGGGACGCGCAACTACCCAGCCCTGGGAACCCCGCGCGG---CGGGCGCG 531
 Db 535 ACTTATGTGTGGATGGCACTTACCATCAGCCCTGGGAACCCAGCGAGGCTACGGGGCGCT 594
 QY 532 GTGGGCGCGGAGCCTATCGGAGGTGAGCGGAGGATCTCTGGGCGCTGCGCACTGAGAG 591
 Db 595 GTCGGTGAAGGTGCTTACCGGAGGTGAGCGGAGGACCCCGCGCGTGGCGTGGAG 654
 QY 592 GCGCTGGTGAAGGTGCTTACCGGAGGTGAGCGGAGGATCTCTGGGCGCTGCGCACTGAG 651
 Db 655 GCGCTGGTGAAGGTGCTTACCGGAGGTGAGCGGAGGATCTCTGGGCGCTGCGCACTGAG 714
 QY 652 GTCATGTACACAGCCCTTCAAGCGCGCTCGGAGATTCGGTGGACTACGTGCGAGTCTGTGAC 711
 Db 715 GTCATGTACACAGCTTTCAGGCACTGGCGGACTCCGTGGACTAGTGCAGGTCTGTGAC 774
 QY 712 TCGGACACAGAGTGGGACCCCATGCGACTGCTGAGCTCGTGGGGTACTGGACAGGAC 771
 Db 775 TCAGACACAGACTAGACCCCATGCGACTGCTGGAGCTTGTGCGAGTGTGGATGAAGAC 834
 QY 772 CCCCGGTAGGGGCTGTGGTGGGAGCTGCGGATCCCTTAACCTCTGAGCTCCCTGGGTC 831
 Db 835 CCCCGGTAGGGGCTGTGGAGGGGATGTGAGGATCCCTTAACCTCTGAGCTCCCTGGGTC 894
 QY 832 AGCTTCTTAAGCAGCTCGGATATCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGC 891
 Db 895 AGCTTCTTGAGCAGCTTTCGATATCTGGGTAGCCTTCAATGTGGAGCAGCTTGTGAGAGC 954
 QY 892 TACTTCCACTGTGTATCTTCGATCAGCGGTCTCTAGGCTCTATAGGAAATAACCTCTTG 951
 Db 955 TACTTCCACTGTGTCTCTGATCAGTGGTCTCTGGGCTCTATACAGAAAATACTCTCTG 1014
 QY 952 CAGCAGTCTTCTGGGCTGTGTACACAGAGATTCCTGGGTACCCACTGCTACTTTTGGG 1011
 Db 1015 CAGCAGTCTTCTGGAGGCGCTGTGTACACCAAAAGTTCTTGGGACCCACTGTCACATTTGG 1074
 QY 1012 GATGACGGGACCTCACCACCGCATGCTCAGCATGGGTTATGCTACCAAGTA 1064
 Db 1075 GATGACGGGACCTCACCACCGCATGCTTAGCATGGGCTATGCTACCAAGTA 1127

RESULT 10

AAZ10862

ID AAZ10862 standard; DNA; 2102 BP.

XX AAZ10862;

XX 19-OCT-1999 (first entry)

XX Hyaluronate synthase coding sequence.
DE Hyaluronate synthase; promoter; mouse; cell growth inhibitor; ds.
KW Mus sp.
XX Location/Qualifiers
FH 49..1800
FT /*tag= a
XX
XX JP11196875-A.
XX 27-JUL-1999.
XX
XX 14-JAN-1998; 98JP-00006191.
XX 14-JAN-1998; 98JP-00006191.
XX (SEKG) SEIKAGAKU KOGYO CO LTD.
XX WPI; 1999-496653/42.
DR P-PSDB; AAY32503.
XX New promoter DNA of hyaluronate synthase - used to specifically express
PT gene with cell growth inhibiting activity.
XX
XX Example 1; Page 8-10; 13pp; Japanese.
XX This sequence represents the mouse hyaluronate synthase coding sequence,
CC and was used to isolate the hyaluronate synthase promoter of the
CC invention. The promoter can be used for specifically expressing a gene
CC having cell growth inhibiting activity
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 73.9%; Score 786.6; DB 2; Length 2102;
Best Local Similarity 84.9%; Pred. No. 1.2e-130;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;
QY 10 CAGGACGGGCCCAAGCCCACTCTCGACGCCCGCCGCTGCTCGCGCTCGCCGAGGGTG 69
DB 55 CAGGACATGCGCAAGCCCTCAGAGGCGCGCTGCTGCTGCGCTGCGCCAGCGAGCA 114
QY 70 CTGACATGCGCTTCTGCGCTGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTG 129
DB 115 CTCAGATCATCTTTGCGCTGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTG 174
QY 130 GTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 189
DB 175 GTTCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 234
QY 190 TCAGCGCACCTGCTGCGCGCAGAGCTCTTTGCGGTACCTGGAGCACCGCGCGGTGCG 249
DB 235 AGCGCACACCTAGTGCGCACAGAGCTCTTGGCTTACCTGGAGCACCGAGGCTGCG 294
QY 250 GCGGCGCG-----GCGCGCGCTGAGATGAGCGCACCGCGCGCGAGTGGCG 294
DB 295 GCTGCGCGCGCTCTTGGCGAAGGGCGCCCTGGATGCGGCGCACTGCGCAGAGGTGCA 354
QY 295 CTGACCATCTCGCTTACCGAGGAGCACCGCGGTACCTGCGCGAGTGGCTGCGCGCG 354
DB 355 CTCACCATCTGAGCTTACCGAGGATCCCGCTTACCTGCGCGAGTGGCTGAGCTTGG 414
QY 355 CGCGCGCTGCTGTACCGCGCGCGCGCTGCGCGCTTCTCATGCTGGTGGATGGCAACCG 414
DB 415 CGCGCTTGTGCTACCGCGCACAGAGGTTACCGCGCTCATGCTGGTGGAGCGCAACGC 474
QY 415 GCGGAGGACCTTACATGGTGCATGTTTCCGCGAGGCTTCTGCTGAGAGGACCGCGCC 474
DB 475 GCTGAGGATCTGTACATGGTGCATGTTTCCGAGAGTCTTCCGCGAGTGGAGCGCGCC 534
QY 475 ACGTACGTGTGGAGCGGCACTTACCACCGCCCTGGGAACCGCGCGCG---CGGCGCG 531

DB 535 ACTTATGTGTGGATGGCACTACCATCAGCCCTGGGAACCAAGCGGAGCTACGGGCGCT 594
QY 532 GTGGCGCGCGAGGCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAG 591
DB 595 GTGCGTGAAGGTGCTTACCGGGAGGTGGAGCGGAGGACCCCGGCGGTGGCGGTGGAG 654
QY 592 CGGCTGGTGAAGACTCGCAGGTGCGTGGCGGCGAGCGCTGGGCGGCGGAGCGCGAG 651
DB 655 CGGCTGGTGAAGAACACGCGAGGTGCGTGGCGGTGCGGCTGGGCGGCGGAGCGTGA 714
QY 652 GTCATGTACACAGAGCTTCAAGCGCTCGGAGATTCGGTGGAGTACGTCAGGTCTGTGAC 711
DB 715 GTCATGTACACAGCTTCAAGGCACTGGCGACTCCGTGGAGTACGTCAGGTCTGTGAC 774
QY 712 TCGGACACAAAGTTGGACCCCATCGCACTGTCGAGCTGTCGGGTTACTGGACGAGGAC 771
DB 775 TCAGACACAGACTAGACCCCATGGCACTGTCGAGCTTGTGCGAGTGTGGATGAAGAC 834
QY 772 CCGCGGTAGGGCTGTTGGTGGGACGTGGCGGATCCTTAACCTCTGGACTCCTGGGTC 831
DB 835 CCGCGGTAGGGCTGTTGGAGGGGATGTGAGGATCCTTAACCTCTGGACTCCTGGGTC 894
QY 832 AGCTTCTTAAGCAGCTGCGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTTCAGAGC 891
DB 895 AGCTTCTTAAGCAGCTTTCGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTTCAGAGC 954
QY 892 TACTTCCACTGTGATCTTGCATCAGCGGCTCTTAGGCTTATAGGAATAAAGCTTTTG 951
DB 955 TACTTCCACTGTGCTGCTGCTCAGTGGTCTCTGCGTCTATACAGAAACAATCTCCTG 1014
QY 952 CAGCAGTTCTTGGAGGCTGTACACAGAGTTCCTGGGTACCCACTGCTACTTTGGG 1011
DB 1015 CAGCAGTTCTTGGAGGCTGTGTACACCAAAAGTCTCTGGGCAACCACTGACATTTGG 1074
QY 1012 GATCAGCGGACCTCACCAACCGCATGCTCAGCATGGTATATGTACCAGTA 1064
DB 1075 GATCAGACGGCACCTCACCAACCGCATGCTTACGATGGCTATGTACCAAGTA 1127
RESULT 11
AAZ88199
ID AAZ88199 standard; cdna to mRNA; 2102 BP.
XX AAZ88199;
XX
DT 27-APR-2000 (first entry)
XX Mouse hyaluronate synthase modified protein HAS1 encoding cdna.
XX Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; ds.
XX Mus sp.
XX Location/Qualifiers
FH 49..1800
FT /*tag= a
FT /product= "HAS1"
FT /note= "modified hyaluronate synthase protein"
XX JP2000004886-A.
XX
PD 11-JAN-2000.
XX
XX 24-JUN-1998; 98JP-00193788.
XX
XX 24-JUN-1998; 98JP-00193788.
XX (SEKG) SEIKAGAKU KOGYO CO LTD.
XX WPI; 2000-140125/13.
DR P-PSDB; AAY68491.

XX A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.

PS Claim 18; Page 12-14; 30pp; Japanese.

XX The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. The present sequence encodes mouse HAS1

XX SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;

Query Match 73.9%; Score 786.6; DB 3; Length 2102;
Best Local Similarity 84.9%; Pred. No. 1.2e-130;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;

QY	10	CAGGACGGCCCAAGCCCACTCTCTCAGCCGCGCTCTCGGCTCGCCGGAGGGTG	69
DB	55	CAGGACATCCMAAGCCCTCAGAGGAGCGCGTCTCTCTGCGCTGCGCAGGCGAGA	114
QY	70	CTGACCATCGCCTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAAGCCCGGG	129
DB	115	CTCAGCATCATCTTGGCTGCTCATCTGGGCTCATGACCTGGGCTTAAGCCCGGG	174
QY	130	GTGCGCTGGCTTCGATCGCTACCGGCTCTTGGCTCTTGGCTCTTGGCTCTTGGCT	189
DB	175	GTTCCTCTGGCTTCAGATCGCTATGAGCTCTTGGCTCTTGGCTCTTGGCTCTTGGCT	234
QY	190	TCAGCGCACTGGTGGCGCAGAGCTCTTGGCTACCTGGCTGAGCAGCGCGGGTGGCGG	249
DB	235	AGCGCACACCTAGTGGCAGAGCTCTTGGCTACCTGGCTGAGCAGCGCGGGTGGCGG	294
QY	250	CGCGCGCG-----GGGCGCGTGGATGAGCAGCAGCGCGCGAGTGTGGCG	294
DB	295	GCTGGCGCGCTCTTGGCGAGAGGCGCGCTGGATGGGCTACGACGAGCGTGGA	354
QY	295	CTGACCATCTCGCTTACAGAGAGACCCCGCTACCTGGCTGAGTGTGGCTGGCTGG	354
DB	355	CTCAGCATCTCAGCTTACAGAGAGATCCCGCTTACCTGGCTGAGTGTGGCTGGCT	414
QY	355	CGCGCGCTGCTACCGCGCGCGCTGGCGCTCTCATGCTGGTGGTGGTGGTGGTGG	414
DB	415	CGCGCTTGTGTACCGCGCACAGAGGTTACGCGTGTGTGTGTGTGTGTGTGTGTGT	474
QY	415	GCCGAGGACCTCTACATGTCGACATGTTCCGCGAGGTCTTCTGCTGAGGAGACCCG	474
DB	475	GCTGAGGATCTGATATGTTGAGCATGTTCCGAGAGTCTTCTGCGATGAGGAGCCG	534
QY	475	ACGTACGTGTGGAGCGGCAACTTACCAAGCGCTTGGGAAACCGCGCGCG-----	531
DB	535	ACTTATGTGTGGATGGCAACTTACCATCAGCGCTTGGGAAACCGCGCGCGCT	594
QY	532	GTGGCGCGGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAG	591
DB	595	GTCGGTGAAGGTGGCTTACCGGAGGTGGAGCGGAGGAGCCCGCGCGCTTGGCGT	654
QY	592	GCGCTGTGTGAGACTTCGAGGTGCTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGG	651
DB	655	GCGCTGTGTGAGACTTCGAGGTGCTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGG	714
QY	652	GTCATGTATACAGCTTCAAGCGCTTCGAGATTCGGTGGTACGTCAGCTGTGTGAC	711
DB	715	GTCATGTATACAGCTTCAAGCGCTTCGAGATTCGGTGGTACGTCAGCTGTGTGAC	774
QY	712	TCGAGACCAAGTGGAGCCCATGGCACTGTGGAGCTGTGGAGCTGTGGAGTGTGG	771
DB	775	TCAGACCAAGACTAGACCCCATGGCACTGTGGAGCTGTGGAGTGTGGAGTGTGG	834

QY	772	CCCCGGGTAGGGGCTGTGGTGGGACCTGCGGATCTTAAACCCCTCTGGACTCTGGGTC	831
DB	835	CCCCGGGTAGGGGCTGTGGTGGGAGGATCTGAGGATCTTAAACCCCTCTGGACTCTGGGTC	894
QY	832	AGCTTCCTAAGCAGCGCTCGATCTAGGCTAGCTTCAATGTGGAGCGGCTTGTGAGAC	891
DB	895	AGCTTCCTTGGAGCAGCTCTTGGATCTAGGCTAGCTTCAATGTGGAGCGGCTTGTGAGAC	954
QY	892	TACTTCCACTGTGTATCTCTGATCAGCGCTCTCTAGGCTTATATAGGAATAACCTCTTG	951
DB	955	TACTTCCACTGTGTCTCTGATCAGGCTCTCTGAGTCTCTGAGTCTATACAGAAACATCTCTG	1014
QY	952	CAGCAGTTTCTTGGGCTCTGATACCAACGAGTTCCTGGGTACCACTGATCTTTGGG	1011
DB	1015	CAGCAGTTTCTTGGGCTCTGATACCAACGAGTTCCTGGGTACCACTGATCTTTGGG	1074
QY	1012	GATGACCGGACCTCACCACCGATGCTAGCATGGTATGCTACCAAGTA	1064
DB	1075	GATGACCGGACCTCACCACCGATGCTAGCATGGTATGCTACCAAGTA	1127

RESULT 12

AAA39987
ID AAA39987 standard; cDNA; 2102 BP.

XX AAA39987;

XX AC AC (first entry)

DT 19-OCT-2000

XX Murine HAS1 cDNA.

DE Murine; HAS1; hyaluronate synthase; marker gene; intracellular loop;
KW gene targeting; ss.

XX Mus sp.

Key Location/Qualifiers
CDS 49..1800
FT /*tag= a
FT /product= "HAS1"

XX JP2000116382-A.

XX 25-APR-2000.

XX 13-OCT-1998; 98JP-00291201.

XX 13-OCT-1998; 98JP-00291201.

XX (SEBK) SEIKAGAKU KOGYO CO LTD.

XX WPI; 2000-369404/32.

XX P-PSDB; AAB09948.

XX DNA used for gene targeting of the hyaluronate synthase gene (HAS1).
XX Example; Page 7-9; 14pp; Japanese.

XX This invention describes a novel DNA which contains an heterologous DNA
CC (I) to be introduced to chromosomal DNA of a host cell, a first and
CC second homologous region DNA (II) connected respectively to the 5' and 3'
CC sides of the introduced DNA, and a negative marker gene (III) expressible
CC in the host cell. (I) contains a positive marker gene expressible in the
CC host cell. (I), (II) and the region encoding the intracellular loop of
CC the hyaluronate synthase (HAS) 1 protein from chromosomal DNA is
CC constituted so as to cause a homologous recombination. The DNA can be
CC used for gene targeting of the HAS1 gene. This sequence encodes the
CC murine HAS1 protein described in the method of the invention

XX SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;

Query Match 73.9%; Score 786.6; DB 3; Length 2102;

Best Local Similarity 84.9%; Pred. No. 1.2e-130;		Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;	
Qy	10	CAGACGCGCCCAAGCCCACTCTGTCAGCCGCGCGCTGCTCGGCTGCGCCGAGGGTG	69
Db	55	CAGGACATGCCAAGCCCTCAGAGCAGCGGTTGCTGCTCTGCGCTGCGCCAGCGAGCA	114
Qy	70	CTGACCATCGCTTTCGCCCTGCTCATCTTGGGCTCATGACCTGGGCTTACGCGCCGGG	129
Db	115	CTCAGATCATCTTTGGCCCTGCTCATCTTGGGCTCATGACCTGGGCTTACGCGCAGGC	174
Qy	130	GTGCGCTGGCTCGATCGCTAGCGCTCTGCGCTTTCGCGCTTACGCGGCTTCTCTT	189
Db	175	GTTCCTCTGGCTTCAGATCGCTATGACTCTCTGGCTTTGGCTCTATGGGCAATCTCTC	234
Qy	190	TCAGCGCACCTGGTGGCGCAGAGCCTCTTTCGCTACCTTGGAGCACCGCGGGTGGCGCG	249
Db	235	AGCGCACACCTAGTGGCACAGAGCTCTTTCGCTTACCTTGGAGCACCGAAGGTTGGCAGG	294
Qy	250	CGCGCGCG-----GGGCGCGTGGATGAGCAGCACCGCGCGCAGTGTGGCG	294
Db	295	GCTGCGCGCGCTCTCTTGGCGAAGGGGCGCCCTGGATGCGGCACTGACAGCGGTGGCA	354
Qy	295	CTGACCATCTCGCTTACGAGGAGACCCCGCTACCTGCGCAGTGCCTGCGCTCGGCC	354
Db	355	CTCACCATCTCAGCCTACAAAGAGATCCCGCTTACCTGGCGCAGTGTGACCTCGCG	414
Qy	355	CGCGCCCTGCTGTATCCCGCGCGCGCTGCGCGCTCTCATGCTGGTGGATGGCAACCGC	414
Db	415	CGCGCTTGTGTACCCGACACAGAGTTACGCGTGTCTCATGTTGGTGGACGGCAACCGC	474
Qy	415	GCGGAGGACCTTACATGGTGACATGTTTCGCGAGGTTCTTCTGTGACGAGAACCCCGCC	474
Db	475	GCTGAGGATCTGTATATGTTGACATGTTTCGAGAAGTCTTTCGCGCATGAGGACCCCGCC	534
Qy	475	ACGTACGTGTGGAGCGCACTTACCACAGCCCTGGGAACCCCGCGCGG---CGGGCGG	531
Db	535	ACTTATGTGGGATGGCACTTACCATCAGCCCTGGGAACAGCGGAGGCTACGGGCGCT	594
Qy	532	GTGGCGCGCGAGCCTATCGGAGGTGGAGCGAGGATCTCTGGCGGCTGCGCAGTGGAG	591
Db	595	GTCGTTGAAGTGTCTTACCGGAGGTGGAGCGGAGGAGCCCGGCGGTTGGCGTGGAG	654
Qy	592	GCGCTGTGAGGACTCTCGCAGGTGCTGTGCGGCGACGCTGGGCGGCGCAAGCGCGAG	651
Db	655	GCGCTGTGTGAGAACACGCGAGGTGCTGTGCGTGGCTCAGCGTTGGGCGGCAACGTTGAG	714
Qy	652	GTCATGTACAGACCTTCAAGCGCTCGAGATTTCGGTGGACTACGTGCAGGTCTGTGAC	711
Db	715	GTCATGTACAGACTTTTCAAGGCACTGGCGCACTCCGTGGACTACGTGCAGGTCTGTGAC	774
Qy	712	TCGGACACAAAGTTTGGACCCCATGTCACCTGCTGGAGCTCTGTCGGGTTACTGGACGAGAC	771
Db	775	TCAGACACAGACTAGACCCCATGSCACTGCTGGAGCTTGTGGAGTGTGGATGAAGAC	834
Qy	772	CCCCGGGTAGGGCTGTTGTGGGGAAGCTGGGGAATCTTAACCTCTTGGACTCTCGGGTC	831
Db	835	CCCCGGGTAGGGCTGTTGTGGGGAAGTGTGGGATCTTAAACCTCTTGGACTCTCGGGTC	894
Qy	832	AGCTTCTTAAAGCCTCGATACCTGAGTGGGTAGCGCTTCAATGTGGAGCGGCTTGTACAGAGC	891
Db	895	AGCTTCTTGAAGCAGTCTTTCGATACCTGAGTGGGTAGCGCTTCAATGTGGAAACGAGCTTGTACGAGC	954
Qy	892	TACTTCCACTGTGTATCTTCGATCAGCGGTCTCTTAGGCCTATATAGGAATAACCTCTTG	951
Db	955	TACTTCCACTGTGTCTCTGATCAGTGTCTCTGGGTCTATACAGAAACAATCTCTCTG	1014
Qy	952	CAGCAGTTTCTTGGGCTGTGTACAAACAGAGTTCTTGGGTACCCACTGTACTTTTGGG	1011
Db	1015	CAGCAGTTTCTTGGGAGGCTGTGTACAAACAGAGTTCTTGGGACCCACTGTACCAATTTGGG	1074
Qy	1012	GATGACCGGCACTTACCAACCGCATGCTCAGCATGGGTATGTCTACCAAGTA 1064	

Db	1075	GATGACAGGCACCTCACCAACCGAATCTTAGCATGGCTATGTCTACCAAGTA	1127
RESULT 13			
ID	ADL13693/C		
XX	ADL13693 standard; DNA; 231222 BP.		
AC	ADL13693;		
XX			
DT	06-MAY-2004 (first entry)		
XX			
DE	Osteoarthritis-associated polymorphic nucleotide #225.		
XX			
KW	ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;		
KW	joint space narrowing; osteophyte development; joint pain;		
KW	osteoarthritis; SNP; single nucleotide polymorphism.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2003054166-A2.		
XX			
PD	03-JUL-2003.		
XX			
PF	19-DEC-2002; 2002WO-US041225.		
XX			
PR	20-DEC-2001; 2001US-0342603P.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
PI	Jones KA, Schafer A;		
XX			
DR	WPI; 2003-559141/52.		
XX			
PT	Determining susceptibility of an individual to joint space narrowing,		
PT	osteoarthritis development and/or joint pain comprises identifying whether		
PT	the individual has at least one polymorphism in a polynucleotide encoding		
XX	a protein.		
PS	Disclosure; SEQ ID NO 225; 297pp; English.		
XX			
CC	The invention relates to a method of determining susceptibility of an		
CC	individual to joint space narrowing and/or osteophyte development and/or		
CC	joint pain comprising identifying whether the individual has at least one		
CC	polymorphism in a polynucleotide encoding at least one of the protein		
CC	listed in the specification. The methods, composition and agent are		
CC	useful for modulating the susceptibility of an individual to joint space		
CC	narrowing and/or osteophyte development and/or joint pain that is		
CC	associated with a disease, preferably osteoarthritis. The cell line and		
CC	the non-human animal are useful for screening for an agent for diagnosing		
CC	an individual having susceptibility to joint space narrowing and/or		
CC	osteoarthritis development and/or joint pain. This sequence corresponds to		
CC	the polynucleotide encoding a protein listed in the specification. (Note:		
CC	The sequence data for this patent did not form part of the printed		
CC	specification but was obtained in electronic format directly from WIPO at		
CC	ftp.wipo.int/pub/published_pct_sequences).		
XX			
SQ	Sequence 231222 BP; 64006 A; 52087 C; 53478 G; 61650 T; 0 U; 1 Other;		
Query Match 65.5%; Score 697.2; DB 10; Length 231222;			
Best Local Similarity 99.6%; Pred. No. 8.4e-115;			
Matches 699; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	7	CAGCAGACGCGCCCAAGCCCACTCTGTCAGCCGCGCGCTGCTCGGCTGCGCGGAGG	66
Db	149122	CAGCAGACGCGCCCAAGCCCACTCTGTCAGCCGCGCGCTGCTCGGCTGCGCGGAGG	149063
Qy	67	GTGCTGACCATCGCTTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCCGCC	126
Db	149062	GTGCTGACCATCGCTTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCCGCC	149003
Qy	127	GGGGTGGCGCTGGCTTCGATCGCTACGGCTCTCGGCTTTCGGGCTTACGGGGCTTTC	186
Db	149002	GGGGTGGCGCTGGCTTCGATCGCTACGGCTCTCGGCTTTCGGGCTTTCGGGCTTTC	148943

QY 187 CTTTCAGCGCACCTGTTGGCGCAGAGCCTCTTCGGGTACCTTGAGCACCAGCGGGTGGCG 246
Db 148942 CTTTCAGCGCACCTGTTGGCGCAGAGCCTCTTCGGGTACCTTGAGCACCAGCGGGTGGCG 148883
QY 247 GCGCGCGCGCGCGCGCTGATGATGAGCACCAGCGCGAGTGTGGCGTGACATCTCC 306
Db 148892 GCGCGCGCGCGCGCGCGCTGATGATGAGCACCAGCGCGAGTGTGGCGTGACATCTCC 148823
QY 307 GCTTACCAGGAGACCCGCGTACCTTGCGCAGTGCCTGGCGTCCGCGCGCGCTGCTG 366
Db 148822 GCTTACCAGGAGACCCGCGTACCTTGCGCAGTGCCTGGCGTCCGCGCGCGCTGCTG 148763
QY 367 TACC CGCGCGCGCTGCGGCTCTCATGTGTGGATGGCAACCGCGCGGAGGACCTC 426
Db 148762 TACC CGCGCGCGCTGCGGCTCTCATGTGTGGATGGCAACCGCGCGGAGGACCTC 148703
QY 427 TACATGTTGACATGTTCCGCGAGTCTTCGCTGACGAGGACCCGCGCACATGACGTGG 486
Db 148702 TACATGTTGACATGTTCCGCGAGTCTTCGCTGACGAGGACCCGCGCACATGACGTGG 148643
QY 487 GAGCGCAATACACAGCCCTGGGAACCCGCGCGCGCGCGTGGCGCGCGAGCC 546
Db 148642 GAGCGCAATACACAGCCCTGGGAACCCGCGCGCGCGCGTGGCGCGCGAGCC 148583
QY 547 TATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGGTGAGGACT 606
Db 148523 TATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGGTGAGGACT 148523
QY 607 CGAGGTGCGTGTGCGCGCAGCGCTGGGGCGGCAAGCGCGAGGTGATGACAGCC 666
Db 148522 CGAGGTGCGTGTGCGCGCAGCGCTGGGGCGGCAAGCGCGAGGTGATGACAGCC 148463
QY 667 TTCAGGCGCTCGGAGATTCGCTGACGCTACGTCGAGTCTGT 708
Db 148462 TTCAGGCGCTCGGAGATTCGCTGACGCTACGTCGAGTCTGT 148421

RESULT 14
ID ABQ27384
XX ABQ27384 standard; DNA; 662 BP.
XX AC ABQ27384;
XX DT 12-JUL-2002 (first entry)
XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 13975.
XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.
XX OS Homo sapiens.
XX PN WO200218632-A2.
XX PD 07-MAR-2002.
XX PF 01-SEP-2001; 2001WO-EP010074.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PR 05-SEP-2000; 2000DE-01044543.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX DR WPI; 2002-371829/40.
XX PT Determining the degree of cytosine methylation in genomic DNA, useful for
XX PT diagnosis and prognosis, comprises selective hybridization of amplicons
XX PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 662 BP; 80 A; 80 C; 247 G; 255 T; 0 U; 0 Other;
Query Match 39.6%; Score 422; DB 6; Length 662;
Best Local Similarity 77.3%; Pred. No. 6.4e-66;
Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY 34 GCAGCCCGCGCTGCTCCGGCTGCGCCCGGAGGGTGTGACCATCGCCTTCGCCCTGCTC 93
Db 1 GTAGTTTCTGCTGTTTTCGGTTTGTTCGAGGGTGTGATTATCGTTTTCGTTTGTGTT 60
QY 94 ATCTCGGCGCTCATGACCTGGCGCTACGCCCGCGGGGTGCGCGCTCCGATCGCTAC 153
Db 61 ATTTGGGGTTTATGATTGGGTTTACGTCGCGGGTGTGCTGTTGGTTTTCGATCGTTAC 120
QY 154 GGCTCTCTGGGCTTCGGGCTCTACGGGCGCTTCCTTTTCAGCGCACCTCGTGGCGCAGAGC 213
Db 121 GGTTTTGTGTTTTCGGTTTTCGGGTTTTCGGGTTTTCGGGTTTTCGGGTTTTCGGGTT 180
QY 214 CTCTTCGCGTACCTGGAGCACCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 273
Db 181 TTTTTCGCGTATTTGGAGTATCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 274 GCGACCGCGCGCGAGTGTGGCGCTGACCATCTCCGCTACGAGGAGGACCCCGCGTACCTG 333
Db 241 GTTATCGCGCGTAGTGTGGCGTTGATTATTTTCGTTTATTAGGAGGATTCGCGTATTG 300
QY 334 CGCCAGTGCCTGGCGTCCGCCCGCGCGCTGCTGTACCCGCGCGCGCGCGCGCGCGCTC 393
Db 301 GGTAGTGTTCGGGCTTCGTTCCGCTTGTGTGTTTTCGCGCGCGCGGTTCGCGGTTTT 360
QY 394 ATGTTGTGGATGGCAACCGCGCGGAGGACCTCTACATGTTGTCGATGTTTCGCGAGGTC 453
Db 361 ATGTTGTGGATGGTAACTCGCTCGAGGATTTTATATGTTGATGTTTCGCGAGGTT 420
QY 454 TTCGCTGACAGGAGACCCCGCGCACGTA CGTGTGGAGCGGCAACTACCAAGCCCTCGGAA 513
Db 421 TTCGTTGACGAGGATTTTCGTTACGTA CGTGTGGAGCGGTAATATTATTAGTTTGGGAA 480
QY 514 CCCGCGCGCGCGCGGTTCGGCGCGGAGGACCTATCGGAGGTGGAGCGGAGGATCCT 573
Db 481 TTCGCGCGCGCGCGGTTCGGCGCGGAGGATTCGAGGATTCGAGGAGGAGGATTTT 540
QY 574 GGGCGGCTGCGCAGTGGAGCGCGCTGCTGAGGAGTCTCGAGTGTGCTGCTGCGCAGCGC 633
Db 541 GGGCGGTTGTTAGTGGAGGCGTTCGTTAGGATTCGTTAGTGTGCTGCTGCGTAGCGT 600
QY 634 TGGGGCGGCAAGCGCGAGGTCATGTACACAGCTTCAAGGCGCTCGGAGATTCGGTGGAC 693
Db 601 TGGGGCGGTAAGCGCGAGGTTATGTATATAGTTTTCGAGGCGTTTCGAGAGATTCGGTGGAT 660

Query Match:
Best Local
Matches 1

Qy	70	CTGACCATCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTACGCCCGGG	129
Db	118	CTGACCATCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTACGCCCGGG	177
Qy	130	GTGCGCTGGCTCGATCGCTACGGCTCTGGGCTTGGGCTTACGGGGCTTCTT	189
Db	178	GTGCGCTGGCTCGATCGCTACGGCTCTGGGCTTGGGCTTACGGGGCTTCTT	237
Qy	190	TCAGCGCACTGGTGGCGAGAGCTCTTCGCGTACCTGGAGCAACCGCGGTGGCGCG	249
Db	238	TCAGCGCACTGGTGGCGAGAGCTCTTCGCGTACCTGGAGCAACCGCGGTGGCGCG	297
Qy	250	GCGCGCGGGGCGCTGGATGACGACCGCGCGAGTGTGGGCTGACCATCTCGCC	309
Db	298	GCGCGCGGGGCGCTGGATGACGACCGCGCGAGTGTGGGCTGACCATCTCGCC	357
Qy	310	TACAGAGGACCCCGCTACTCTGGCCAGTGCCTGGGCTCGCGCGGCTTCTGTAC	369
Db	358	TACAGAGGACCCCGCTACTCTGGCCAGTGCCTGGGCTCGCGCGGCTTCTGTAC	417
Qy	370	CCGCGCGCGCTGGCTGCTCATGTGTGTGATGGCAACCGCGCGAGGACCTCTAC	429
Db	418	CCGCGCGCGCTGGCTGCTCATGTGTGTGATGGCAACCGCGCGAGGACCTCTAC	477
Qy	430	ATGTCGACATGTTCCGCGAGTCTTCTGCTGACGAGGACCGCGCATGCTGTGGAC	489
Db	478	ATGTCGACATGTTCCGCGAGTCTTCTGCTGACGAGGACCGCGCATGCTGTGGAC	537
Qy	490	GGCAACTACCAACAGCCCTGGGAAACCGCGCGCGGCTGGGCGCGGAGCTTAT	549
Db	538	GGCAACTACCAACAGCCCTGGGAAACCGCGCGCGGCTGGGCGCGGAGCTTAT	597
Qy	550	CGGAGGTGGAGCGGAGATCTCTGGCGCTGGCGAGTGGAGGCGCTGGTGAAGACTGC	609
Db	598	CGGAGGTGGAGCGGAGATCTCTGGCGCTGGCGAGTGGAGGCGCTGGTGAAGACTGC	657
Qy	610	AGTTCGCTGCTGGTGGCGAGGCTTGGGCGCGGCGGAGGCTCATGTACAGGCTTC	669
Db	658	AGTTCGCTGCTGGTGGCGAGGCTTGGGCGCGGCGGAGGCTCATGTACAGGCTTC	717
Qy	670	AAGCGCTCGAGATTCGGTGGACTACCTGAGGCTCTGACTCGGACACAAAGGTTGGAC	729
Db	718	AAGCGCTCGAGATTCGGTGGACTACCTGAGGCTCTGACTCGGACACAAAGGTTGGAC	777
Qy	730	CCATGGCACTGCTGGAGCTCGTGGGTACTGGACGAGGACCCCGGGTAGGGCTGTT	789
Db	778	CCATGGCACTGCTGGAGCTCGTGGGTACTGGACGAGGACCCCGGGTAGGGCTGTT	837
Qy	790	GGTGGGACGTCGGATCTTAAACCTCTGAGCTCTGGGTCAGTTCTTAAGCAGCTG	849
Db	838	GGTGGGACGTCGGATCTTAAACCTCTGAGCTCTGGGTCAGTTCTTAAGCAGCTG	897
Qy	850	CGATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGACTCTTCCACTGTGTATCC	909
Db	898	CGATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGACTCTTCCACTGTGTATCC	957
Qy	910	TGATCTAGCGCTCTTAGGCTATATAGGAATAACCTTTGACGAGATTTCTTGAAGCC	969
Db	958	TGATCTAGCGCTCTTAGGCTATATAGGAATAACCTTTGACGAGATTTCTTGAAGCC	1017
Qy	970	TGTTACACAGAGATTCCTGGTACCACTGTACTTTTGGGATGACCGGCACTTACC	1029
Db	1018	TGTTACACAGAGATTCCTGGTACCACTGTACTTTTGGGATGACCGGCACTTACC	1077
Qy	1030	AACGCACTGCTCAGCATGGGTATGCTACCAAGTA 1064	
Db	1078	AACGCACTGCTCAGCATGGGTATGCTACCAAGTA 1112	

RESULT 2
AY421569
LOCUS

1728 bp DNA linear GSS 17-DEC-2003

DEFINITION Homo sapiens HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY421569
VERSION AY421569.1 GI:39748428
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1728)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1728)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1728
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1728
/gene="HAS1"
/locus_tag="HCM7601"
ORIGIN
Query Match 98.8%; Score 1051.8; DB 9; Length 1728;
Best Local Similarity 99.8%; Pred. No. 8.8e-210;
Matches 1053; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 10 CAGGACGCGCCCAAGCCCACTCTCGAGCGCGCGCTGCTCCGGCTTGGCCCGGAGGTG 69
Db 1 CAGGACGCGCCCAAGCCCACTCTCGAGCTTGGCGCTTGGCCCTTGGCCCGGAGGTG 60
Qy 70 CTGACCATCGCTTGGCCCTGCTCATCTCGGGCTCATGACCTGGGCTACGCCCGCGG 129
Db 61 CTGACCATCGCTTGGCCCTGCTCATCTCGGGCTCATGACCTGGGCTACGCCCGCGG 120
Qy 130 GTGCGCTGGCTCCGATCGCTACGGCTCTCGGCTTGGGCTTGGGCTTCTTCTT 189
Db 121 GTGCGCTGGCTCCGATCGCTACGGCTCTCGGCTTGGGCTTCTTCTTCTTCTT 180
Qy 190 TCAGCGACCTGTGGCGCAGAGCTCTTCGCTACCTGGAGCACCGCGGCTGGCGCG 249
Db 181 TCAGCGACCTGTGGCGCAGAGCTCTTCGCTACCTGGAGCACCGCGGCTGGCGCG 240
Qy 250 GCGCGCGGGGCGGCTGGATGAGCACCGCGCGAGTGTGGCGCTGACCATCTCGCC 309
Db 241 GCGCGCGGGGCGGCTGGATGAGCACCGCGCGAGTGTGGCGCTGACCATCTCGCC 300
Qy 310 TACAGAGGACCCCGGTACCTGCGCAGTGTGGCTTGGGCTTGGGCTTCTTCTTCTT 369
Db 301 TACAGAGGACCCCGGTACCTGCGCAGTGTGGCTTGGGCTTGGGCTTCTTCTTCTT 360
Qy 370 CCGCGCGCGGCTTGGCGCTTCTCATGTGTGTGATGGCAACCGCGCGGAGGACTCTAC 429
Db 361 CCGCGCGCGGCTTGGCGCTTCTCATGTGTGTGATGGCAACCGCGCGGAGGACTCTAC 420
Qy 430 ATGTCGACATGTTCCGCGAGGCTTCTGCTGACGAGGACCCCGGCACTGCTGTGGAC 489
Db 421 ATGTCGACATGTTCCGCGAGGCTTCTGCTGACGAGGACCCCGGCACTGCTGTGGAC 480


```
Db      835 GNGGGGAGCTGCGGATCTTAACCTCTGGACTCTGGTCAGTTTCTTAGCAACCTG 894
Qy      850 CGAT-ACTGGGTAGCTTCAAT-GTGGAGCGGGCTTG 884
Db      895 CGAANAAGTGGGTAGCTTCAATGGTGAACGGGCTTG 931

RESULT 4
AL545447      965 bp      mRNA      linear      EST 25-MAR-2004
LOCUS      AL545447 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION      clone CS0DI027YJ02 5-PRIME, mRNA sequence.
ACCESSION      AL545447
VERSION      AL545447.3 GI:45745929
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31267282.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DI027DE01QP1&c=9502.r.

FEATURES
source
1..965
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CS0DI027YJ02"
    /issue_type="PLACENTA COT 25-NORMALIZED"
    /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
    /note="1st strand cDNA was primed with a NotI-oligo(dT)
    primer. Five prime end enriched, double-strand cDNA was
    digested with Not I and cloned into the Not I and EcoR V
    sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      74.7%; Score 795.6; DB 1; Length 965;
Best Local Similarity 97.4%; Pred. No. 3.6e-156;
Matches 851; Conservative 9; Mismatches 8; Indels 6; Gaps 5;

Qy      10 CAGGACGGCCCAAGCCACTCTTCAGCCGCGCGCTCTCGGCTGGCCCGCGAGGGTG 69
Db      58 CAGGACGGCCCAAGCCACTCTTCAGCCGCGCGCTCTCGGCTGGCCCGGA-GGTG 116
Qy      70 CTGACCATCGCTTCGCCCTGCTCATCTCTGGGCTCATGACTGGGCTACGCCCGGG 129
Db      117 CTGACCATCGCTTCGCCCTGCTCATCTCTGGGCTCATGACTGGGCTACGCCCGGG 176
Qy      130 GTGCGCTGGGCTTCGATCGCTACGGCTCTCGGCTTCGGGCTCTACGGGGCTTCCTT 189
Db      177 GTGCGCTGGGCTTCGATCGCTACGGCTCTCGGCTTCGGGCTCTACGGGGCTTCCTT 236
Qy      190 TCAGCGCACTGTGGCGCAGAGCTCTTCGGTACTTGAGACACCGCGGGTGGCGCG 249
Db      237 TCAGCGCACTGTGGCGCAGAGCTCTTCGGTACTTGAGACACCGCGGGTGGCGCG 296
Qy      250 GCGGCGGGGCGCGCTGGTGGAGCCACCGCGCGAGTGTGGCGCTGACCATCTCCGCC 309

297 GCGCGCGGGGCGCTGGATGCAGCCACCGCGCAGTGTGGCGCTGACCATCTCCGCC 356
310 TACCAGAGGACCCCGGCTACCTCGCGCAGTGCCTTGGGCTCCGCCCGGCCCTTGTGTAC 369
357 TACCAGGAGGACCCCGGCTACCTCGCGCAGTGCCTTGGGCTCCGCCCGGCCCTTGTGTAC 416
370 CCGCGCGCGGCTGCGCTCCTCATGTGTGTGATGCAACCGCGCGGAGAGCTCTTAC 429
417 CCGCGCGCGGCTGCGGCTCCTCATGTGTGTGATGCAACCGCGCGGAGAGCTCTTAC 476
430 ATGTCGACATGTTTCCCGGAGGCTTCGCTGACAGGACCCCGCCACGTCAGTGTGGAC 489
477 ATGTCGACATGTTTCCCGGAGGCTTCGCTGACAGGACCCCGCCACGTCAGTGTGGAC 536
490 GGCAACTACCAAGCCCTTGGGAAACCGCGCGCGCGCGGCTGGGCGCGGAGCCTAT 549
537 GGCAACTACCAAGCCCTTGGGAAACCGCGCGCGCGCGGCTGGGCGCGGAGCCTAT 596
550 CCGGAGGTGGAGGCGGAGGATCCTGGGCGGCTGGCAGTGGAGGCGCTGGTGAGGACTCGC 609
597 CCGGAGGTGGAGGCGGAGGATCCTGGGCGGCTGGCAGTGGAGGCGCTGGTGAGGACTCGC 656
610 AGTGTGCTGTGCTGGGCGCAGCGCTGGGCGGCGCAAGCGCGAGGTCACTGACACAGCCTTC 669
657 AGTGTGCTGTGCTGGGCGCAGCGCTGGGCGGCGCAAGCGCGAGGTCACTGACACAGCCTTC 716
670 AAGCGCTCGGAGATTCGCTGGGACTACGTGAGGCTCTGTGACTCGGACACAGGTTGGAC 729
717 AAGCGCTCGGAGATTCGCTGGGACTACGTGAGGCTCTGTGACTCGGACACAGGTTGGAC 776
730 CCCATGCACTGCTGGAGCTGCTGCGGCTGCTGACGAGGACCCCGGGTAGGGGCTGTT 789
777 CCCATGCACTGCTGGAGCTGCTGCGGCTGCTGACGAGGACCCCGGGTAGGGGCTGTT 836
790 GGTGGGACGTGGCGGATCCTTAACCTCTGAGCTCTCGGCTGAGCTTCCTTAAGCAGCGCTG 849
837 GGTGGGACGTGGCGGAT-CYTAMCCTCTGAGCTCTCTKGTAR--CTTCTAAGCAG-CTG 892
850 CGATACCTGGGTAGCTTCAATGTGGAGCGGGCTT 883
893 CGATACCTGGGTAG-CTTAAATGTGGAGCGGGTTK 925

RESULT 5
AY421571      1746 bp      DNA      linear      GSS 17-DEC-2003
LOCUS      AY421571 Mus musculus HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AY421571
VERSION      AY421571.1 GI:39748430
KEYWORDS      GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1746)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1746)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
```

```
them based on alignment.
FEATURES             Location/Qualifiers
  source             1..1746
                    /organism="Mus musculus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10090"
  gene               <1..>1746
                    /gene="HAS1"
                    /locus_tag="HCM7601"

ORIGIN
Query Match       74.0%; Score 788.2; DB 9; Length 1746;
Best Local Similarity 85.0%; Pred. No. 1.3e-154;
Matches 912; Conservative 0; Mismatches 143; Indels 18; Gaps 2;

QY 10 CAGGACGGCCCAAGCCACTCTCTGAGCCCGCCCTGCTCGGCTGCGCCGGAGGTG 69
   |||||
Db 1 CAGGACATGCCAAGCCCTCAGAGGAGCGGCTTGCTGCTGCGCTGCGCCAGCGGCA 60

QY 70 CTGACCATGCGCTTCGCCCTCTCATCTCGGCTCATGACCTGGGCTTACGCCCGCGG 129
   |||||
Db 61 CTCAGATCATCTTTGCCCTGCTCATCTCGGCTCATGACCTGGGCTTACGCCCGAGC 120

QY 130 GTGCGCTGCGCTTCGATCGCTACGGCTCTTGGCTTCGGCTCTACGGGCTTCCTT 189
   |||||
Db 121 GTTCTCTGCTTCAGATCGCTATGAGCTCTCTGGCTTTGGCTCTATGGGCAATCTCT 180

QY 190 TCAGCGCACCTCGTGGCGAGAGCTCTTCGGGTACCTGGAGCACCGGGGTGGCGG 249
   |||||
Db 181 AGCGCACACTAGTGCGCAGAGCTCTTCGGTTACCTGGAGCACCGAAGGTGGCAGC 240

QY 250 GCGGCGCG-----GGGCGCGCTGATGACGACCGCGCGCAGTGTGGCG 294
   |||||
Db 241 GCTGCGCGCGCTCTCTTGCGAAGGGGCCCTGGATGCGGCACATGACGAGCGTGCA 300

QY 295 CTGACCATCTCGCTTACAGAGAGACCCCGGTAACCTGGCGCACTGCTGCGCTCGCG 354
   |||||
Db 301 CTCACCATCTCAGCCCTACCAAGAGATCCCGCTTACCTGGCGCAGTGTGACCTCCG 360

QY 355 CGCGCCCTGCTGTACCGCGCGCGCTGCGCGCTCCTCATGCTGGTGGTGGCAACGC 414
   |||||
Db 361 CGCGCTTGTGCTGTACCGGCACAGAGTTACGCGTGCTCATGTTGGTGGACGGCAACGC 420

QY 415 GCCGAGGACCTTACATGCTGACATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCC 474
   |||||
Db 421 GCTGAGGATCTGTACATGCTGACATGTTCCGAGAGTCTTCGCGATGAGGACCCCGC 480

QY 475 AGTACGTGTGGGACGGCAATACCAACAGCCCTTGGGAACCCCGCGCGG---CGGCGCG 531
   |||||
Db 481 ACTTATGTGTGGATGGCAACTPACCATCAGCCCTGGGAACCAAGCGGAGGCTACGGCGCT 540

QY 532 GTGGCGCGGAGCTATCGGAGGTGAGCGGAGGATCTCTGGCGGCTGCGAGTGGAG 591
   |||||
Db 541 GTGCGTGAAGTGCTTACCGGAGGTGAGCGGAGGACCCCGCGCGTGTGCGGTGGAG 600

QY 592 GCGCTGTGAGGACTCGCAGTGTGCTGCGGCGCAGCGCTGGGCGGCAAGCGCAG 651
   |||||
Db 601 GCGCTGTGTAGAACACGCAAGTGTGCTGCGTGGCTCAGCGTTGGGCGGCAAGCGTGA 660

QY 652 GTCATGTACAGCCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGAGGTCTGTGAC 711
   |||||
Db 661 GTCATGTACAGCTTTTCAAGSCACTGGGCGACTCCGTGGACTACGTGCGAGTCTGTGAC 720

QY 712 TCGGACACAGGTTGAGCCCATGCACTGCTGGAGCTCGTGGGTTACTGGACGAGAC 771
   |||||
Db 721 TCAGACACAAGACTAGACCCCATGCACTGTGTGAGCTTGTGCGAGTGTGTGATGAAGAC 780

QY 772 CCGCGGTAGGGCTGTGTTGGGAGCGTGGGATCTTAACTCTGACTCTCTGGGTC 831
   |||||
Db 781 CCGCGGTAGGGCTGTGTTGGAGGAGTGTGAGATCTTAACTCTCTGACTCTCTGGGTC 840

QY 832 AGCTTCTTAAGCAGCTCGATATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGC 891
   |||||
```


Db 778 TCGCTGACGAGGACCCCGCCAGTACGTGT-GGACGGCAACTACCAACGAGCCCTGGGAAA 720
Qy 514 CCCGGCGGGCGGGCGGGTGGCGCCGAGCCTATCGGAGGTGAGCGGAGGATCCT 573
Db 719 CCNCNGCNGCGGCGGGTGGCCCGGAGCCTATCGGGAGGTGNAGNCAAGGATCCT 660
Qy 574 GGGCGGCTGGCAGTGGAGCGCTGTGTGAGGACTCGCAGGTGGTGGTGGCGCGAGCGC 633
Db 659 GNCGGCTGGCAGTGGAGCGCTGTGTGAGGACTCGCAGGTGGTGGTGGCGAGCGC 600
Qy 634 TGGCGCGCAAGCGCGAGGTCAATACAGCCTTCAAGCGCTCGGAGATTCGGTGGAC 693
Db 599 TGGCGCGCAAGCGCGAGGTCAATACAGCCTTCAAGCGCTCGGAGATTCGGTGGAC 540
Qy 694 TACGTGCGAGGTGTGACTCGGACCAAGGTGGAGCCCATGGCACTCTCGAGGCTCGT 753
Db 539 TACGTGCGAGGTGTGACTCGGACCAAGGTGGAGCCCATGGCACTCTCGAGGCTCGT 480
Qy 754 CGGTGACTGGAGGAGCCCGCGGTAGGGCTGTGTGGGAGCTGCGGATCCTTAAC 813
Db 479 CGGTGACTGGAGGAGCCCGCGGTAGGGCTGTGTGGGAGCTGCGGATCCTTAAC 420
Qy 814 CCTCGGACTCCTGGGTGAGCTTCTTAAGCAGCCTCGGATCTGGGTAGCCTTCAATGTG 873
Db 419 CCTCGGACTCCTGGGTGAGCTTCTTAAGCAGCCTCGGATCTGGGTAGCCTTCAATGTG 360
Qy 874 GAGCGGCTTGTGAGGACTTCTCACTGTGTATCTGTCATCAGCGGTCTCTAGGCTTA 933
Db 359 GAGCGGCTTGTGAGGACTTCTCACTGTGTATCTGTCATCAGCGGTCTCTAGGCTTA 300
Qy 934 TATAGGAATAACTCTTCGAGCAGTTCTTGAGGCTGGTACAACCAAGTTCCTGGGT 993
Db 239 TATAGGAATAACTCTTCGAGCAGTTCTTGAGGCTGGTACAACCAAGTTCCTGGGT 240
Qy 994 ACCCACTGACTTTGGGATGACCGCACTTCAACCAACCGCATGCTCAGCATGGTTAT 1053
Db 239 ACCCACTGACTTTGGGATGACCGCACTTCAACCAACCGCATGCTCAGCATGGTTAT 180
Qy 1054 GCTACCAAGTA 1064
Db 179 GCTACCAAGTA 169

RESULT 7
AK053726
LOCUS AK053726 3542 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30302P16 product:hyaluronan synthase1, full insert sequence.
ACCESSION AK053726
VERSION AK053726.1 GI:26343670
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
2
20499374
PUBMED 11042159
REFERENCE 10349636
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED 10349636
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, I., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED 11076861
4

THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 3542)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL:http://genome.gsc.riken.jp/

URL:http://fantom.gsc.riken.jp/.

Location/Qualifiers

1. 3542

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM.DB:EI30302P16"

/db_xref="taxon:10090"

/clones="EI30302P16"

/tissue_type="eyeball"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="0 day neonate"

61. 456

/note="unnamed protein product; hyaluronan synthase1

(MGDI:MG1:106590, GB|NM_008215, evidence: BLASTN, 99%, match=762)

putative"

/codon_start=1

/protein_id="BAC35492.1"

/db_xref="GI-26343671"

/translation="MRQDMFKPSEARCCSGLARRALTITFALLILGLMTWAYAGVP

LASDRGLLAFGLYGAFLSAHLVAQSLFAYLEHRRVAAARRSLAKGLDATTARSA

LTISAYQEDPAYLRQCLTSARALLYPTR"

CDS

ORIGIN

Query Match 45.0%; Score 479.4; DB 3; Length 3542;
Best Local Similarity 81.6%; Pred. No. 5.6e-90;
Matches 598; Conservative 0; Mismatches 116; Indels 19; Gaps 3;
QY 10 CAGGACGGCCCAAGCCACTCTCGACGCGCCCGCTGCTCGGCTTGGCCCGGAGGGTG 69
DB 67 CAGGACGGCCCAAGCCCTCAGAGGCGCGCTTCTGCTTGGCTTGGCCCGGAGCA 126
QY 70 CTGACCATCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTACGCCCGCGG 129
DB 127 CTCAGATCATCTTGGCTGCTCATCTGGGCTCATGACCTGGGCTTACGCCCGAGGC 186
QY 130 GTGCGCTGGCTTCCGATCGCTACCGCTCTGGCTTGGCTTACCGGCTTACCGGCTTCTT 189
DB 187 GTTCTCTGGCTTCAGATCGCTATGACCTCTGGCTTGGCTTCTATGGGCAATTCCTC 246
QY 190 TCAGCGCACTTGGGCGAGAGCTTCTCGGTACCTTGGACCGGCGGCTTGGCGG 249
DB 247 AGCGCACACCTAGTGGCACAGAGCTCTTGGCTTACCTGGAGCACCGAAGGGTGGCAGCG 306
QY 250 GCGGCGCG-----GGGCGCGCTGGATGACGACCGCGCGCGAGTGTGGCG 294
DB 307 GCTGGCGCGCTCTTGGCGAGGGGCGCTTGGATGGGCGCACTGCGAGCGGTGGCA 366
QY 295 CTGACCATCTCGGCTTACGAGGAGGACCGCGGTACCTTGGCGAGTGTGGCTTGGCGCC 354
DB 367 CTCACCATCTCAGCTTACCAAGAGGATCCGCTTACCTTGGCGAGTGTGGCTTGGCGG 426
QY 355 CGCGCTTGTGTACCGCGCGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 414
DB 427 CGCGCTTGTGTGTACCGCGCACAGAGGTAACGCGTGTCTCATGCTGTGGAGCGGCAACGCG 486
QY 415 GCCGAGGACCTTACATGCTGCATGTTTCGCGAGGCTTTCGCTGACGAGGACCGCGCC 474
DB 487 GCTGAGGATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 546
QY 475 ACCTATGCTGGATGGCACTTACCATCAGCCCTGGGAACCAAGCGGCTTACGCGGCT 606
QY 532 GTGGCGCGGAGCTTACGAGGAGTGGAGGAGGATCTTGGCGGCTTGGCGGCTTGGCGG 591
DB 607 GTGCGTGAAGGTGCTTACCGGAGGTGGAGGCGGAGGACCGCGGCGGCTTGGCGG 666
QY 592 GCGCTGTGTAGGAC--TCGACGCTGCTGTGGTGGCGAGCGCTTGGCGGCGGAGCGGCA 650
DB 667 GCGCTGTGTAGGACAGCGGAGTGGCTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 726
QY 651 GGTATGTATACAGCTTCAAGGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGA 710
DB 727 GGTATGTATACAGCTTCAAGGCACTGGGCGACTCCGCTGAGTACGTGCAAGTGTGA 786
QY 711 CTCGACACAAAGG 723
DB 787 TGGGTTTCCAGG 799

RESULT 8
BM544718
LOCUS
DEFINITION AGENCOURT 6494603 NIH_MGC_124 Homo sapiens cdna clone IMAGE:5727416
5', mRNA sequence.
ACCESSION BM544718
VERSION BM544718.1 GI:18776197
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1084)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12721 row: i column: 15
High quality sequence start: 5
High quality sequence stop: 689.
Location/Qualifiers
1. 1084
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5727416"
/issue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV (destroyed); Site 2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."

FEATURES

source

ORIGIN

Query Match 41.5%; Score 442.4; DB 4; Length 1084;
Best Local Similarity 99.8%; Pred. No. 3e-82; Mismatches 0; Indels 0; Gaps 0;
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 621 CGTGGCGAGCGCTGGGCGCGCAAGCGAGGTCTGTGACTCGGACACAAAGTTGGACCCCATGGCACT 680
DB 20 CGTGGCGAGCGCTGGGCGCGTAAGCGCGAGGTCTGTGACTCGGACACAAAGTTGGACCCCATGGCACT 79
QY 681 AGATTGGTGGACTACGTGCGAGGTCTGTGACTCGGACACAAAGTTGGACCCCATGGCACT 740
DB 80 AGATTGGTGGACTACGTGCGAGGTCTGTGACTCGGACACAAAGTTGGACCCCATGGCACT 139
QY 741 GCTGAGCTCGTGGGCTTACTGGAGGAGCCCGGGTAGGGGTGGTGGGGAGCT 800
DB 140 GCTGAGCTCGTGGGCTTACTGGAGGAGCCCGGGTAGGGGTGGTGGGGAGCT 199
QY 801 GCGGATCCTTAACCCCTCTGGACTCTGGGCTCAGCTTCTTAAGCAGCCCTGCATATCTGGGT 860
DB 200 GCGGATCCTTAACCCCTCTGGACTCTGGGCTCAGCTTCTTAAGCAGCCCTGCATATCTGGGT 259
QY 861 AGCTTCAATGTGAGCGGGCTTGTGAGGACTATTCACCTGTGTATCTTCATCATCAGCGG 920
DB 260 AGCTTCAATGTGAGCGGGCTTGTGAGGACTATTCACCTGTGTATCTTCATCATCAGCGG 319
QY 921 TCCTCTAGGCTATATAGGATACCTCTTGCAGCAGTTTCTTGCAGGCTGTGTACCAACA 980
DB 320 TCCTCTAGGCTATATAGGATACCTCTTGCAGCAGTTTCTTGCAGGCTGTGTACCAACA 379
QY 981 GAAGTTCCTGGTACCACCATGTACTTTTGGGGATGACCGGACCTCACCACCGCATGCT 1040
DB 380 GAAGTTCCTGGTACCACCATGTACTTTTGGGGATGACCGGACCTCACCACCGCATGCT 439
QY 1041 CAGCATGGTGTATGCTACCAAGTA 1064
DB 440 CAGCATGGTGTATGCTACCAAGTA 463

RESULT 9

BG707429
LOCUS
DEFINITION 602672848F1 NIH_MGC_96 Homo sapiens cdna clone IMAGE:4795643 5',

```

mRNA sequence.
BG707429      1  GI:13983769
EST.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 551)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10678 row: p column: 12
High quality sequence stop: 551.
FEATURES             source
    Location/Qualifiers
        1..551
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4795643"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_96"
            /note="Organ: brain; Vector: pBluescriptR (modified
            pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
            (gtcgag); Oligo-dT primed using primer
            5'-TTTTTTTTTTTTTTVN-3', size-selected for average
            insert size 2.3 kb and normalized to ROT 5. This is a
            primary library enriched for full-length clones and
            constructed using the Cap-trapper method (Carninci, in
            preparation). Library constructed by M. Brownstein
            (NIMH/NHGRI, National Institutes of Health). Note: this is
            a NIH_MGC Library."
ORIGIN
Query Match      40.2%; Score 428.4; DB 4; Length 551;
Best Local Similarity 99.8%; Pred. No. 2.5e-79;
Matches 429; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

635  GGGGGCGACGCGAGGTCATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGACT 694
      |||||||
Db    6  GGGGGCGCAAGCGGAGGTCATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGACT 65

695  ACGTGCAGGTCGTGACTCGGACACAAGGTTGGACCCCATGGCACTGCTGGAGCTCGTGC 754
      |||||||
Db    66  ACGTGCAGGTCGTGACTCGGACACAAGGTTGGACCCCATGGCACTGCTGGAGCTCGTGC 125

755  GGGTACTGGACGAGACCCCGGTAGGGCTGTGTTGGTGGGACGTGGGATCCTTAACC 814
      |||||||
Db    126  GGGTACTGGACGAGACCCCGGTAGGGCTGTGTTGGTGGGACGTGGGATCCTTAACC 185

815  CTCTGGACTCTGGGTGAGCTTCTTAAGACGCTTGCATGCTGATCTGCTGAGTACCTCAATCGG 874
      |||||||
Db    186  CTCTGGACTCTGGGTGAGCTTCTTAAGACGCTTGCATGCTGATCTGCTGAGTACCTCAATCGG 245

875  AGCGGGCTTGTGACAGCTACTTCCACTGTGTATCTGTCATCAGCGGCTCTTAGGCCCTAT 934
      |||||||
Db    246  AGCGGGCTTGTGACAGCTACTTCCACTGTGTATCTGTCATCAGCGGCTCTTAGGCCCTAT 305

935  ATAGGAATAACCTCTTGACGAGCTTTCTTAGGCGCTCGGTACAACAGAGTTCCTGGGTA 994
      |||||||
Db    306  ATAGGAATAACCTCTTGACGAGCTTTCTTAGGCGCTCGGTACAACAGAGTTCCTGGGTA 365

mRNA sequence.
BG707429      1  GI:13983769
EST.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 551)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10678 row: p column: 12
High quality sequence stop: 551.
FEATURES             source
    Location/Qualifiers
        1..551
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4795643"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_96"
            /note="Organ: brain; Vector: pBluescriptR (modified
            pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
            (gtcgag); Oligo-dT primed using primer
            5'-TTTTTTTTTTTTTTVN-3', size-selected for average
            insert size 2.3 kb and normalized to ROT 5. This is a
            primary library enriched for full-length clones and
            constructed using the Cap-trapper method (Carninci, in
            preparation). Library constructed by M. Brownstein
            (NIMH/NHGRI, National Institutes of Health). Note: this is
            a NIH_MGC Library."
ORIGIN
Query Match      40.2%; Score 428.4; DB 4; Length 551;
Best Local Similarity 99.8%; Pred. No. 2.5e-79;
Matches 429; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

635  GGGGGCGACGCGAGGTCATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGACT 694
      |||||||
Db    6  GGGGGCGCAAGCGGAGGTCATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGACT 65

695  ACGTGCAGGTCGTGACTCGGACACAAGGTTGGACCCCATGGCACTGCTGGAGCTCGTGC 754
      |||||||
Db    66  ACGTGCAGGTCGTGACTCGGACACAAGGTTGGACCCCATGGCACTGCTGGAGCTCGTGC 125

755  GGGTACTGGACGAGACCCCGGTAGGGCTGTGTTGGTGGGACGTGGGATCCTTAACC 814
      |||||||
Db    126  GGGTACTGGACGAGACCCCGGTAGGGCTGTGTTGGTGGGACGTGGGATCCTTAACC 185

815  CTCTGGACTCTGGGTGAGCTTCTTAAGACGCTTGCATGCTGATCTGCTGAGTACCTCAATCGG 874
      |||||||
Db    186  CTCTGGACTCTGGGTGAGCTTCTTAAGACGCTTGCATGCTGATCTGCTGAGTACCTCAATCGG 245

875  AGCGGGCTTGTGACAGCTACTTCCACTGTGTATCTGTCATCAGCGGCTCTTAGGCCCTAT 934
      |||||||
Db    246  AGCGGGCTTGTGACAGCTACTTCCACTGTGTATCTGTCATCAGCGGCTCTTAGGCCCTAT 305

935  ATAGGAATAACCTCTTGACGAGCTTTCTTAGGCGCTCGGTACAACAGAGTTCCTGGGTA 994
      |||||||
Db    306  ATAGGAATAACCTCTTGACGAGCTTTCTTAGGCGCTCGGTACAACAGAGTTCCTGGGTA 365

```

```

QY 815 CTCTGGACTCCTGGGTGAGCTTCTTAAGCAGCTCGGATCTGGGTAGGCTTCAATGTGG 874
Db 186 CTCTGGACTCCTGGGTGAGCTTCTTAAGCAGCTCGGATCTGGGTAGGCTTCAATGTGG 245
QY 875 AGCGGGCTTGTGAGGCTTCTTCACTGTGTATCTTCAATGAGCGGTCTCTAGGCTTAT 934
Db 246 AGCGGGCTTGTGAGGCTTCTTCACTGTGTATCTTCAATGAGCGGTCTCTAGGCTTAT 305
QY 935 ATAGGAATAAAGCTTCTTGGGATGACCGGACCTCACCACCGCATGCTCAGCATGGGTATG 994
Db 306 ATAGGAATAAAGCTTCTTGGGATGACCGGACCTCACCACCGCATGCTCAGCATGGGTATG 365
QY 995 CCCACTGTACTTTTGGGATGACCGGACCTCACCACCGCATGCTCAGCATGGGTATG 1054
Db 366 CCCACTGTACTTTTGGGATGACCGGACCTCACCACCGCATGCTCAGCATGGGTATG 425
QY 1055 CTACCAAGTA 1064
Db 426 CTACCAAGTA 435

RESULT 11
BX327795 610 bp mRNA linear EST 08-APR-2004
DEFINITION BX327795 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1027YJ02 5-PRIME, mRNA sequence.
ACCESSION BX327795
VERSION BX327795
KEYWORDS BX327795.2 GI:46277978
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 610)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30340686.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-Oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0BAG028ZH06_CS02640_1&c=9502.r

FEATURES
source
Location/Qualifiers
1..610
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1027YJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 37.1%; Score 395.4; DB 5; Length 610;
Best Local Similarity 98.3%; Pred. No. 2.1e-72;
Matches 399; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 659 ACACAGCCTTCAAGCGCTCGGAGATTCGGTGGAGTACGTGAGCTGTGTGACTCGGACA 718
Db 1 ACATGCTCTTAAGGCGCTCGGAGATTCGGTGGAGTACGTGAGCTGTGTGACTCGGACA 60

```

```

QY 719 CAAGTTGGACCCCATGCACTGCTGGAGCTCGTGGGGTACTGGACGAGACCCCCCGG 778
Db 61 CAAGTTGGACCCCATGCACTGCTGGAGCTCGTGGGGTACTGGACGAGACCCCCCGG 120
QY 779 TAGGGGCTGTTGGTGGGAGACGTGGGGATCCCTTAACCCCTCTGGAATCTCTGGGTGAGCTTCC 838
Db 121 TAGGGGCTGTTGGTGGGAGACGTGGGGATCCCTTAACCCCTCTGGAATCTCTGGGTGAGCTTCC 180
QY 839 TAAGCAGCTCGGATCTGAGTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCC 898
Db 181 TAAGCAGCTCGGATCTGAGTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCC 240
QY 899 ACTGTGTATCTGCACTCAGCGGTCTCTAGGCGCTATATAGGAATAAACCCTTTGCAGCAGT 958
Db 241 ACTGTGTATCTGCACTCAGCGGTCTCTAGGCGCTATATAGGAATAAACCCTTTGCAGCAGT 300
QY 959 TTCTTGAGGCTGTGTACACGAGAGTTCCTGGGTACCCACTGTACTTTGGGATGACC 1018
Db 301 TTCTTGAGGCTGTGTACACGAGAGTTCCTGGGTACCCACTGTACTTTGGGATGACC 360
QY 1019 GGCACCTCACCACCGCATGCTCAGCATGGGTATGCTACCAAGTA 1064
Db 361 GGCACCTCACCACCGCATGCTCAGCATGGGTATGCTACCAAGTA 406

RESULT 12
BX327795 782 bp mRNA linear EST 25-SEP-2001
DEFINITION 603025961P1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196399 5',
mRNA sequence.
ACCESSION BI753116
VERSION BI753116.1 GI:15744694
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 782)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11491 Row: j Column: 16
High quality sequence stop: 780.
Location/Qualifiers
1..782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5196399"
/clone_lib="NIH_MGC_114"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

FEATURES
source
Location/Qualifiers
1..782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5196399"
/clone_lib="NIH_MGC_114"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match 35.7%; Score 380.4; DB 4; Length 782;

```



```

QY 415 GCGGAGGACCTTACATGTCGACATGTTCCGCGAGTCTTCGCTGACGAGGACCCGCC 474
    |||||
Db 487 GCTGAGGATCTGTACATGTCGACATGTTCCGAGAAGTCTTCGCCGATGAGGACCCGCC 546
    |||||
QY 475 AGCTAGTGTGGAGCGGCACTACACACCGCTGGGAACCCGCCGCCGCCGCCGGTG 534
    |||||
Db 547 ACTTATGTGTGGATGGCACTACCATCAGCCCTGGGACCGGGAAGCTACAGGCGCTG 606
    |||||
QY 535 GCGCGCGGAGCCTATCGGAGGTGGAGCGGAGGATCCTGGGCG 578
    |||||
Db 607 TCGGTCAAGGCGCTACCGGAGGTGAGCGAGACCCCGGGCG 650

RESULT 14
LOCUS AY421570
DEFINITION Pan troglodytes HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION AY421570
VERSION AY421570.1 GI:39748429
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 739)
AUTHORS Clark,A.G., Gnanawong,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 739)
AUTHORS Clark,A.G., Gnanawong,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..739
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>739
/gene="HAS1"
/locus_tag="HCM7601"
ORIGIN
Query Match 29.7%; Score 315.8; DB 9; Length 739;
Best Local Similarity 99.4%; Pred. No. 9.4e-56;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 746 AGCTGTGGGGTACTGACGAGGACCCCGGTTAGGGCTGTTGGTGGGACGTGGCGA 805
    |||||
Db 1 AGCTGTGGGGTACTGACGAGGACCCCGGTTAGGGCTGTTGGTGGGACGTGGCGA 60
    |||||
QY 806 TCCTTAACCTCTGACCTCTGGTTCAGTCTTCTTAAGCAGCTCGGATCTGGGTAGCCT 865
    |||||
Db 61 TCCTTAACCTCTGACCTCTGGTTCAGTCTTCTTAAGCAGCTCGGATCTGGGTAGCCT 120
    |||||
QY 866 TCAATGTGGAGCGGCTTGTTCAGAGCTACTTCCATCTGTATCTTCGTCATCAGCGTCTCTC 925
    |||||
Db 121 TCAATGTGGAGCGGCTTGTTCAGAGCTACTTCCATCTGTATCTTCGTCATCAGCGTCTCTC 180
    |||||
QY 926 TAGGCCTATATAGGAATACCTCTTGGCAGCAGTCTTCTTAGGCCTGGTACACCAAGAGT 985
    |||||

```

```

Db 181 TAGGCCTCTATAGGAATAACCTCTTTCGACGAGTCTTCTTAGGCCTGGTACACCAAGAGT 240
QY 986 TCCTGGGTACCCACCTGTACTTTTGGGGATGACCGCACCTCACCACCGCATGCTCAGCA 1045
    |||||
Db 241 TCCTGGGCACCCACCTGTACTTTTGGGGATGACCGCACCTCACCACCGCATGCTCAGCA 300
    |||||
QY 1046 TGGGTTATGCTTACCAAGTA 1064
    |||||
Db 301 TGGGTTATGCTTACCAAGTA 319
    |||||

RESULT 15
LOCUS AK028582
DEFINITION Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
library, clone:4732404L04 product:similar to DG42III [Brachydanio
rerio], full insert sequence.
ACCESSION AK028582
VERSION AK028582.1 GI:26324525
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 20499374
MEDLINE 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yanamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 20530913
MEDLINE 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hasegawa,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akai,H.,

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:01:16 ; Search time 209.778 Seconds
(without alignments)
8307.054 Million cell updates/sec

Title: US-10-672-399-7
Perfect score: 1085
Sequence: 1 atgagacagcaggacgcgc.....tgggttatgtaccacagtaa 1065

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	99.9	2116	3	US-08-635-552A-1 Sequence 1, Appli
2	1060.8	99.6	2087	4	US-09-949-016-5659 Sequence 5659, Ap
3	1051.6	98.7	2117	3	US-09-155-768-1 Sequence 1, Appli
4	697.2	65.5	14862	4	US-09-949-016-17401 Sequence 17401, A
5	396.4	37.2	601	4	US-09-949-016-201775 Sequence 201775,
6	353	33.1	1662	4	US-09-949-016-3728 Sequence 3728, Ap
7	311.6	28.3	1665	4	US-08-812-008-31 Sequence 31, Appli
8	246.8	23.2	2947	4	US-08-675-499A-1 Sequence 1, Appli
9	246.8	23.2	2947	4	US-08-812-008-1 Sequence 1, Appli
10	230.4	21.6	601	4	US-09-949-016-201780 Sequence 201780,
11	221	20.8	3003	4	US-09-949-016-737 Sequence 737, App
12	167.6	15.7	9871	4	US-09-949-016-15470 Sequence 15470, A
13	137.4	12.9	601	4	US-09-949-016-132199 Sequence 132199,
14	127.8	12.0	32176	4	US-09-949-016-12479 Sequence 12479, A
15	95.6	9.0	235	4	US-08-675-499A-25 Sequence 25, Appl
16	95.6	9.0	235	4	US-08-812-008-25 Sequence 25, Appl
17	92.4	8.7	235	4	US-08-675-499A-26 Sequence 26, Appl
18	92.4	8.7	235	4	US-08-812-008-26 Sequence 26, Appl
19	88.2	8.3	235	4	US-08-675-499A-23 Sequence 23, Appl
20	88.2	8.3	235	4	US-08-812-008-23 Sequence 23, Appl
21	82.8	7.8	6645	4	US-09-902-540-8819 Sequence 8819, Ap
22	82.8	7.8	10210	4	US-09-902-540-938 Sequence 938, App
23	69.2	6.5	1320	4	US-09-724-797-85 Sequence 85, Appl
24	66.6	6.3	3957	4	US-10-237-551-193 Sequence 193, App
25	66.6	6.3	154746	4	US-09-827-688-8 Sequence 8, Appli
26	66.6	6.3	154746	4	US-09-827-688-8 Sequence 8, Appli
27	66.2	6.2	1614	4	US-09-616-289-45 Sequence 45, Appl

c 28	66.2	6.2	12425	4	US-09-616-289-50	Sequence 50, Appl
c 29	64.8	6.1	536165	4	US-09-214-808-1	Sequence 1, Appli
c 30	62	5.8	1292	3	US-08-483-533-37	Sequence 37, Appl
31	62	5.8	1292	3	US-09-283-471A-37	Sequence 37, Appl
32	60.6	5.7	4257	2	US-08-690-473-1	Sequence 1, Appli
33	60.6	5.7	4257	3	US-09-259-821A-1	Sequence 1, Appli
34	60.6	5.7	4257	3	US-08-843-659-1	Sequence 1, Appli
35	60.6	5.7	4257	4	US-09-825-288A-1	Sequence 1, Appli
c 36	60.6	5.7	12001	1	US-08-458-568A-11	Sequence 11, Appl
37	60.2	5.7	2900	1	US-08-034-650-9	Sequence 9, Appli
38	60.2	5.7	2900	1	US-08-449-015-9	Sequence 9, Appli
39	59.6	5.6	1647	4	US-09-902-540-4211	Sequence 4211, Ap
c 40	59.6	5.6	12330	4	US-09-902-540-1209	Sequence 1209, Ap
c 41	59.4	5.6	1926	3	US-09-249-585A-4	Sequence 4, Appli
c 42	59.4	5.6	1931	2	US-09-130-114-2	Sequence 2, Appli
43	59.4	5.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli
44	59.2	5.6	1051	2	US-08-865-273-1	Sequence 1, Appli
45	59.2	5.6	1051	3	US-09-385-174-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-635-552A-1
; Sequence 1, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millicia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
US-08-635-552A-1

Query Match 99.9%; Score 1064; DB 3; Length 2116;
Best Local Similarity 100.0%; Pred. No. 1.1e-203;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACAGCAGGACGCCCAAGCCCACTCCTGACGCCGCGCTGCTCGGCGCTGGCC 60
|||||

```
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGCGCCTGGCC 95
QY 61 CGGAGGGTGTGACCATGCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db 96 CGGAGGGTGTGACCATGCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 155
QY 121 GCGCGCGGGTGGCGCTGCGCTCCGATCGTACGAGCCTCTGGCTTGGCTTCTACGGG 180
Db 156 GCGCGCGGGTGGCGCTGCGCTCCGATCGTACGAGCCTCTGGCTTGGCTTCTACGGG 215
QY 181 GCCTCTCTTTTACGCGACCTGGTGGCGAGCCTCTTGGCTTCTGGAGCACCGCGG 240
Db 216 GCCTCTCTTTTACGCGACCTGGTGGCGAGCCTCTTGGCTTCTGGAGCACCGCGG 275
QY 241 GTGGCGGGCGGGCGGGGCGCTGATGAGCAGCAGCGCGCAGTGTGGCGCTGACC 300
Db 276 GTGGCGGGCGGGCGGGGCGCTGATGAGCAGCAGCGCGCAGTGTGGCGCTGACC 335
QY 301 ATCTCCGCTTACAGGAGACCCCGCTACCTGCGCCAGTGCCTGGCGTCCGCGCGCC 360
Db 336 ATCTCCGCTTACAGGAGACCCCGCTACCTGCGCCAGTGCCTGGCGTCCGCGCGCC 395
QY 361 CTGCTGTATACCGCGCGGGCTGGCGCTCTCATGTGGTGGATGGCAACCGCGCGAG 420
Db 396 CTGCTGTATACCGCGCGGGCTGGCGCTCTCATGTGGTGGATGGCAACCGCGCGAG 455
QY 421 GACCTCTACATGTTGCGAGATTCCTGAGTGTGGCGGGTGGCGAGTGTGGCGCTGACC 480
Db 456 GACCTCTACATGTTGCGAGTGTGGCGGGTGGCGAGTGTGGCGCTGACC 515
QY 481 GTGGGAGCGGCAACTACACAGCCCTGGGAAACCGCGGGCGGGCGGGTGGCGGCC 540
Db 516 GTGGGAGCGGCAACTACACAGCCCTGGGAAACCGCGGGCGGGCGGGTGGCGGCC 575
QY 541 GGAGCCTATCGGAGGTGGAGCGGAGTCTTGGCGGGTGGCGAGTGGAGCGCGTGGTG 600
Db 576 GGAGCCTATCGGAGGTGGAGCGGAGTCTTGGCGGGTGGCGAGTGGAGCGCGTGGTG 635
QY 601 AGGACTCGGAGTGTGGTGGCGAGCGCTGGGCGGGCAAGCGCGAGTGTGATG 660
Db 636 AGGACTCGGAGTGTGGTGGCGAGCGCTGGGCGGGCAAGCGCGAGTGTGATG 695
QY 661 ACAGCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGAGTCTGGACTCGGACACA 720
Db 696 ACAGCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGAGTCTGGACTCGGACACA 755
QY 721 AGTTGGGACCCCATGGCACTGTGGAGCTCGTGGCGGTACTGGAGAGGACCCCGGGTA 780
Db 756 AGTTGGGACCCCATGGCACTGTGGAGCTCGTGGCGGTACTGGAGAGGACCCCGGGTA 815
QY 781 GGGGCTGTGGTGGGAGCGTGGGATCTTTAACCTCTGGACTCTGGGTCAGCTTCTTA 840
Db 816 GGGGCTGTGGTGGGAGCGTGGGATCTTTAACCTCTGGACTCTGGGTCAGCTTCTTA 875
QY 841 AGCAGCTGCGGATCTGGGATGCTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCAC 900
Db 876 AGCAGCTGCGGATCTGGGATGCTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCAC 935
QY 901 TGTGTATCTGATCAGCGGCTCTTAGCGCTATATAGGAATAAATCTTTCAGCAGTTT 960
Db 936 TGTGTATCTGATCAGCGGCTCTTAGCGCTATATAGGAATAAATCTTTCAGCAGTTT 995
QY 961 CTTGAGGCTGTGTAACAACAGAGTTCTTGGGTACCCACTGTGATCTTTTGGGAGTACCGG 1020
Db 996 CTTGAGGCTGTGTAACAACAGAGTTCTTGGGTACCCACTGTGATCTTTTGGGAGTACCGG 1055
QY 1021 CACCTTCAACCAACCGATGCTGAGCATGGTTATGCTACCAAGTA 1064
Db 1056 CACCTTCAACCAACCGATGCTGAGCATGGTTATGCTACCAAGTA 1099
```

RESULT 2

US-09-949-016-5659

```
; Sequence 5659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-5659
```

Query Match 99.6%; Score 1060.8; DB 4; Length 2087;
Best Local Similarity 99.8%; Pred. No. 4.6e-203;
Matches 1062; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGCTTGGCC 60
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGCTTGGCC 95
QY 61 CGGAGGGTGTGACCATGCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db 96 CGGAGGGTGTGACCATGCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 155
QY 121 GCCCGCGGGTGGCGCTGCGCTTCCGATCGTACCGGCTCTTGGCTTCTGCGCTTACCGG 180
Db 156 GCGCGCGGGTGGCGCTGCGCTTCCGATCGTACCGGCTCTTGGCTTCTGCGCTTACCGG 215
QY 181 GCCTCTCTTTTACGCGACCTGGTGGCGAGCCTCTTGGCTTCTGGAGACCGCGGG 240
Db 216 GCCTCTCTTTTACGCGACCTGGTGGCGAGCCTCTTGGCTTCTGGAGACCGCGGG 275
QY 241 GTGGCGGGCGGGCGGGGCGCTGGATGCGAGCAGCGCGCAGTGTGGCGCTGACC 300
Db 276 GTGGCGGGCGGGCGGGGCGCTGGATGCGAGCAGCGCGCAGTGTGGCGCTGACC 335
QY 301 ATCTCCGCTTACAGGAGGACCCCGCGCTACCTGGCGCAGTGCCTGGCGTCCGCGCGCC 360
Db 336 ATCTCCGCTTACAGGAGGACCCCGCGCTACCTGGCGCAGTGCCTGGCGTCCGCGCGCC 395
QY 361 CTGCTGTATACCGCGCGGGCTGGCGCTCTCATGTGGTGGATGGCAACCGCGCGGAG 420
Db 396 CTGCTGTATACCGCGCGGGCTGGCGCTCTCATGTGGTGGATGGCAACCGCGCGGAG 455
QY 421 GACCTCTACATGTTGCGAGATTCCTGAGTGTGGCGGGTGGCGAGTGTGGCGCTGACC 480
Db 456 GACCTCTACATGTTGCGAGTGTGGCGGGTGGCGAGTGTGGCGCTGACC 515
QY 481 GTGGGAGCGGCAACTACACAGCCCTGGGAAACCGCGGGCGGGCGGGTGGCGGCC 540
Db 516 GTGGGAGCGGCAACTACACAGCCCTGGGAAACCGCGGGCGGGCGGGTGGCGGCC 575
QY 541 GGAGCCTATCGGAGGTGGAGCGGAGTCTTGGCGGGTGGCGAGTGGAGCGCGTGGTG 600
Db 576 GGAGCCTATCGGAGGTGGAGCGGAGTCTTGGCGGGTGGCGAGTGGAGCGCGTGGTG 635
QY 601 AGGACTCGGAGTGTGGTGGCGAGCGCTGGGCGGGCAAGCGCGAGTGTGATG 660
Db 636 AGGACTCGGAGTGTGGTGGCGAGCGCTGGGCGGGCAAGCGCGAGTGTGATG 695
QY 661 ACAGCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGAGTCTGGACTCGGACACA 720
```


Db 696 ACAGCCTTCAAGCGGCTCGAGATTGGTGGACTAGTGCAGGTCGTGACTCGAGACA 755
Qy 721 AGGTTGGACCCCATGCACTGCTGGAGCTCGTGGGGTACTGGACGAGGACCCCGGGTA 780
Db 756 AGGTTGGACCCCATGCACTGCTGGAGCTCGTGGGGTACTGGACGAGGACCCCGGGTA 815
Qy 781 GGGGCTGTTGGTGGGACGTCGGGATCCTTAACCTCTGGAATCTCTGGGTCAAGTTCCTA 840
Db 816 GGGGCTGTTGGTGGGATGCGGATCCTTAACCTCTGGAATCTCTGGGTCAAGTTCCTA 875
Qy 841 AGCAGCTGCGATACCTGGGTAGCCTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
Db 876 AGCAGCTGCGATACCTGGGTAGCCTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 935
Qy 901 TGTGTATCTGTCATCAGCGGCTCTTAGGCTTATAGGAATACTCTTGCAGCAGTTT 960
Db 936 TGTGTATCTGTCATCAGCGGCTCTTAGGCTTATAGGAATACTCTTGCAGCAGTTT 995
Qy 961 CTTGAGGCTGTGACAAACAGAGTTCTCTGGGTACCCACTGTTTGGGGATGACCGG 1020
Db 996 CTTGAGGCTGTGACAAACAGAGTTCTCTGGGTACCCACTGTTTGGGGATGACCGG 1055
Qy 1021 CACCTCACCAACCGCATGCTCAGCATGGTGTATGCTACCAAGTA 1064
Db 1056 CACCTCACCAACCGCATGCTCAGCATGGTGTATGCTACCAAGTA 1099

RESULT 3

US-09-155-768-1
; Sequence 1, Application US/09155768A
; Patent No. 6162908
; GENERAL INFORMATION:
; APPLICANT: SEIKAGAKU CORPORATION
; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF
; TITLE OF INVENTION: HUMAN-ORIGIN AND DNA ENCODING THE SAME
; FILE REFERENCE: TOYAM37.001APC
; CURRENT APPLICATION NUMBER: US/09/155,768A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: JP 8-084326
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: JP 8-109663
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (149)...(1777)
US-09-155-768-1

Query Match 98.7%; Score 1051.6; DB 3; Length 2117;

Best Local Similarity 99.6%; Pred. No. 3.2e-201;

Matches 1054; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 CAGCAGGACGCGCCCAAGCCCACTCTGCAGCGCGCGCTGCTCGGCTGCGCGGAGG 66
Db 50 CAGCAGGACGCGCCCAAGCCCACTCTGCAGCGCGCGCTGCTCGGCTGCGCGGAGG 109
Qy 67 GTGCTGACCATCGCCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCCGCC 126
Db 110 GTGCTGACCATCGCCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCCGCC 169
Qy 127 GGGGTGCGGCTGCGCTTCGATCGCTACGGCTCTCTGGGCTTGGGCTTACGGGCTTTC 186
Db 170 GGGGTGCGGCTGCGCTTCGATCGCTACGGCTCTCTGGGCTTGGGCTTACGGGCTTTC 229
Qy 187 CTTTCAGCGCACTGCTGGCGCAGAGCTCTTCGGGTACCTTGGACACCGCGGCTGGCG 246
Db 230 CTTTCAGCGCACTGCTGGCGCAGAGCTCTTCGGGTACCTTGGACACCGCGGCTGGCG 289

RESULT 4

US-09-949-016-17401
; Sequence 17401, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

Qy 247 GCGCGCGCGCGGGGCGCTGATGACGCCACCGCGCAGTGTGGCGCTGACCATCTCC 306
Db 290 GCGCGCGCGCGGGGCGCTGATGACGCCACCGCGCAGTGTGGCGCTGACCATCTCC 349
Qy 307 GCCTACAGGAGGACCCCGCGTACTCTGGCCAGTGCCTTGGCGTCCGCGCGCGCCCTGCTG 366
Db 350 GCCTACAGGAGGACCCCGCGTACTCTGGCCAGTGCCTTGGCGTCCGCGCGCGCCCTGCTG 409
Qy 367 TACCGCGCGCGCGCTGCGGCTCCTCATGTGGTGGATGGCAACCGCGCGCGAGGACCTC 426
Db 410 TACCGCGCGCGCGCTGCGGCTCCTCATGTGGTGGATGGCAACCGCGCGCGAGGACCTC 469
Qy 427 TACATGGTCGATATTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCGCTGCTGAGGACT 486
Db 470 TACATGGTCGATATTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCGCTGCTGAGGACT 529
Qy 487 GACGGCACTACCAACGAGGCTTGGGAACCCCGCGCGCGCGGCTGGGCGCGCGAGCC 546
Db 530 GACGGCACTACCAACGAGGCTTGGGAACCCCGCGCGCGCGGCTGGGCGCGCGAGCC 589
Qy 547 TATCGGAGGTGGAGGCGGAGATCCTCGGCGGCTGGCAGTGGAGGCGCTGCTGAGGACT 606
Db 590 TATCGGAGGTGGAGGCGGAGATCCTCGGCGGCTGGCAGTGGAGGCGCTGCTGAGGACT 649
Qy 607 CCGAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTGATGTACACGCC 666
Db 650 CCGAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTGATGTACACGCC 709
Qy 667 TTCAAGGCGCTCGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACCAAGTTG 726
Db 710 TTCAAGGCGCTCGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACCAAGTTG 769
Qy 727 GACCCCATGGCAGTCTCGAGCTGTGGGGTACTTGGAGGAGGACCCCGGGTAGGGCT 786
Db 770 GACCCCATGGCAGTCTGTGAGCTGTGGGGTACTTGGAGGAGGACCCCGGGTAGGGCT 829
Qy 787 GTTGGTGGGACGTCGCGGATCCTTAACCTCTGGAATCTCTGGGTGAGTCTTCTTAAGCAGC 846
Db 830 GTTGGTGGGATGTGCGGATCCTTAACCTCTGGAATCTCTGGGTGAGTCTTCTTAAGCAGC 889
Qy 847 CTGCGATACCTGGTAGCCTTCAATGTGGAGCGGCTTGTGAGACTTTCACACTGTGTA 906
Db 890 CTGCGATACCTGGTAGCCTTCAATGTGGAGCGGCTTGTGAGACTTTCACACTGTGTA 949
Qy 907 TCCTGCATCAGCGGCTCTTAGGCTATATAGGAATAACCTTTCGACAGTTCCTTTCGAG 966
Db 950 TCCTGCATCAGCGGCTCTTAGGCTATATAGGAATAACCTTTCGACAGTTCCTTTCGAG 1009
Qy 967 GCCTGGTACAAACAGAGTTCCTGGGTACCCACTGTACTTTTGGGGATGACCGGACCTC 1026
Db 1010 GCCTGGTACAAACAGAGTTCCTGGGTACCCACTGTACTTTTGGGGATGACCGGACCTC 1069
Qy 1027 ACCAACCGCATGCTCAGCATGGTGTATGCTACCAAGTA 1064
Db 1070 ACCAACCGCATGCTCAGCATGGTGTATGCTACCAAGTA 1107

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17401
; LENGTH: 14862
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17401

Query Match
Best Local Similarity 65.5%; Score 697.2; DB 4; Length 14862;
Matches 699; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 CAGCAGGAGCGCCCAAGCCCACTCTCGAGCCCGCCGCTGCTCGGCTGCGCCCGGAGG 66
Db 6073 CAGCAGGAGCGCCCAAGCCCACTCTCGAGCCCGCCGCTGCTCGGCTGCGCCCGGAGG 6132

QY 67 GTGCTGACCATCGCTTCGCGCTGCTCATCTGCGGCTCATGACCTGGGCTTACGCCGCC 126
Db 6133 GTGCTGACCATCGCTTCGCGCTGCTCATCTGCGGCTCATGACCTGGGCTTACGCCGCC 6192

QY 127 GGGGTGCGGCTGCGCTCCGATCGCTACGCGCTCTTGGGCTTTCGCGCTTACGCGGCTTC 186
Db 6193 GGGGTGCGGCTGCGCTCCGATCGCTACGCGCTCTTGGGCTTTCGCGCTTACGCGGCTTC 6252

QY 187 CTTTCAGCGCACCTGGTGGCGAGAGCCTTTCGCGTACCTGGAGCACCGGGGGTGGCG 246
Db 6253 CTTTCAGCGCACCTGGTGGCGAGAGCCTTTCGCGTACCTGGAGCACCGGGGGTGGCG 6312

QY 247 GGGGGCGCGGGGGCGCGTGGATGACGACCGCGCGAGTGTGGGCTGACCATCTCC 306
Db 6313 GGGGGCGCGGGGGCGCGTGGATGACGACCGCGCGAGTGTGGGCTGACCATCTCC 6372

QY 307 GCCTACAGGAGGACCCCGCTGCTGCGCAGTGCCTGGCGTCCCGCGCGCCCTGCTG 366
Db 6373 GCCTACAGGAGGACCCCGCTGCTGCGCAGTGCCTGGCGTCCCGCGCGCCCTGCTG 6432

QY 367 TACCGCGCGCGCGCTGGGCTGCTCATGTGTTGGATGGCAACCGCGCGCGAGACCTC 426
Db 6433 TACCGCGCGCGCGCTGGGCTGCTCATGTGTTGGATGGCAACCGCGCGCGAGACCTC 6492

QY 427 TACATGTTGACATGTTCCGCGAGTCTTCGCTGACGAGGACCCCGCCACAGTACGTGTGG 486
Db 6493 TACATGTTGACATGTTCCGCGAGTCTTCGCTGACGAGGACCCCGCCACAGTACGTGTGG 6552

QY 487 GACGCCAACTACCAACAGCCCTTGGAAACCCGCGCGCGCGCGCGGTGGCGCCGAGGCC 546
Db 6553 GACGCCAACTACCAACAGCCCTTGGAAACCCGCGCGCGCGCGCGGTGGCGCCGAGGCC 6612

QY 547 TATCGGAGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGGCGCTGGTGGAGACT 606
Db 6613 TATCGGAGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGGCGCTGGTGGAGACT 6672

QY 607 CGCAGTCCGTGCTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTCTATGTACACAGCC 666
Db 6673 CGCAGTCCGTGCTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTCTATGTACACAGCC 6732

QY 667 TTCAGGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGT 708
Db 6733 TTCAGGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGT 6774

RESULT 5
US-09-949-016-201775/c
; Sequence 201775, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3728
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3728

Query Match
33.1%; Score 353; DB 4; Length 1662;
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201775
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201775

Query Match
Best Local Similarity 37.2%; Score 396.4; DB 4; Length 601;
Matches 397; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 309 CTACGAGGAGGACCCCGGTACCTGCGCCAGTCTGCGCTCGCGCCGCGCCCTGCTGTA 368
Db 601 CTACGAGGAGGACCCCGGTACCTGCGCCAGTCTGCGCTCGCGCCGCGCCCTGCTGTA 542

QY 369 CCCCGCGCGCGCTGCGCTCTCTCATGTGTTGGATGGCAACCGCGCGCGAGGACCTCTA 428
Db 541 CCCCGCGCGCGCTGCGCTCTCTCATGTGTTGGATGGCAACCGCGCGCGAGGACCTCTA 482

QY 429 CATGGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGAACCCCGCCACGTAAGTGTGGGA 488
Db 481 CATGGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGAACCCCGCCACGTAAGTGTGGGA 422

QY 489 CGGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGCTGGCGCCGAGGACCTA 548
Db 421 CGGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGCTGGCGCCGAGGACCTA 362

QY 549 TCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAGGCGCTGTGTGAGGACTCG 608
Db 361 TCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAGGCGCTGTGTGAGGACTCG 302

QY 609 CAGTGGGTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCTATGTACAGGCTT 668
Db 301 YAGTGGGTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCTATGTACAGGCTT 242

QY 669 CAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGT 708
Db 241 CAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGT 202

RESULT 6
US-09-949-016-3728
; Sequence 3728, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3728
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3728

Query Match
33.1%; Score 353; DB 4; Length 1662;
```



```
QY 355 CGCGCCCTGCTGTACCGCGCGCGCTGCGGCTCCTCATGGTGGTGGCAACGC 414
Db 358 CAGCGCATGGCTTTTCAAAC-----CTCAAGGTGGTCATGGTAGTGGCAATCGC 381
QY 415 GCGGAGGACCTTACATGGTGCACATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCC 474
Db 382 CAGGAAGATACCTACATGTTGACATCTTCATGAGGTGCTGGGTGGCACTGAGCAAGCT 441
QY 475 ACGTACGTGTGGAGCGGAACTACCAACAGCCCTGGGAAACCGCGCGCGCGCGCGGTG 534
Db 442 GCGTCTTTGTGTGGCGTAGCAATTTCCA-----TGAGCGCGGTGAAGGA 486
QY 535 GCGCGCGAGCTATCGGAGGTGAGCGCGGAGGATCTTGGCGGCTGGCAGTGGAGCG 594
Db 487 GAGACAGAGCGAGCTCGAGGAAGGCATGAGCGT-----GTGCGAGCT 531
QY 595 CTGCTGAGGACTCGCAGGTGCGTGTGCGCGCAGCGCTGGGCGGCGCAAGCGCAGGTC 654
Db 532 GTGGTGTGGCCAGCACCTTCTCATGTCATGTCAGAAAGTGGGGGGCAAGGTGAGGTC 591
QY 655 ATGTACACAGCCTTCAAGCGCTCGGAGATTGGGTGGAAGTCAAGTCAAGTCTGTGACTCG 714
Db 592 ATGTACACCTGCTTCAAGCGCTTGGCAACTCAGTGGACTACATCCAGGTGTGTGACTCT 651
QY 715 GACACAGGTGGACCCCATGGCACTGCTGGAGCTCGTGGGTAAGTGGACGAGACCCC 774
Db 652 GACACTGTGCTGGACCCAGCTGCAACCATGAGATGCTTCGAGTCTTGGAAAGAGATCCC 711
QY 775 CGGTGAGGGCTGTGGTGGGACGTGGGATCTTAAACCTCTCGACTCTCGGTGAGC 834
Db 712 CAAGTAGAGGTGTGGAGGAGATGTCAAATCTCAACAAGTATGATTCATGATCTCC 771
QY 835 TTCTTAAGCAGCTCGATACCTGAGTACCTTCAATGTGGAGCGGCTTGTCAAGTAC 894
Db 772 TTCTTGAAGGTGTGGAGGATGTCATCAGAAAGTTCCTTAGCAGCAAGTCTTCCAG 831
QY 895 TTCCACTGTATCTCTGATCAGCGCTCTTAGCCCTATATAGAAATACCTCTTGAG 954
Db 832 TTGGCTGTGTGCAATGATGATGAGTGGGCTTTGGGCAATGATCCGCAACAGGCTTCCAG 891
QY 955 CAGTTCTTGGAGGCTGTGACCAAGGATCTTGGGTAAGTCTTGGGAT 1014
Db 892 CAGTTCTTGGAGGATGATGATCAGAAAGTTCCTTAGCAGCAAGTGTGAGCTTTGGGAT 951
QY 1015 GACCGCACCTTACCAACCGAGTCTGAGTCTTGGGTAAGTGTGAGTCTTGGGAT 1064
Db 952 GATCGCACCTTACCAACCGAGTCTGAGTCTTGGGTAAGTGTGAGTCTTGGGAT 1001
```

RESULT 8

```
US-08-675-499A-1
; Sequence 1, Application US/08675499A
; Patent No. 6492150
```

GENERAL INFORMATION:

```
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HVALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
```

COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2947 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-675-499A-1
```

Query Match 23.2%; Score 246.8; DB 4; Length 2947;

Best Local Similarity 57.8%; Pred. No. 2e-40;

Matches 524; Conservative 0; Mismatches 337; Indels 45; Gaps 3;

```
QY 159 CCTGCCCTTCGGCCCTCTACGGGGCTTCTCTTTACGACACCTGTGGCGCAGAGCCTCTT 218
Db 636 CTTCTCTATTTGGACTGTAGCGTGCCTTTTATAGCTCGCATCTCATCTCAAAAGCCTCTT 695
QY 219 CGGTACTCTGAGACACCGCGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 278
Db 696 TGCCTTTTGGAAACACCGGAAATGAAGAAATCCCTTGAACACCCGATTTAAATGAACAA 755
QY 279 CGCGCGCAGTGTGGCGCTGACCATCTCCGCTACAGAGGAGACCCGCGTACCTGCGGCA 338
Db 756 AAG-----GTAGCATCTGTCATCGTGTGATCAAGAGACCTGACTACTTACGAA 809
QY 339 GTGCTCGGCTCGCCCGCGCGCTGTGTATCCCGCGCGCGCGCGCGCGCGCGCTCTCATG 398
Db 810 ATGTTTGAATCTGTGAAAGGCTGACCTACCTG-----GGATTAAAGTCGTGATGGT 863
QY 399 GTGTGATGGCAACCGCGCGCGGACCTTACATGTCGACATGTCGCGAGGTCTTCGC 458
Db 864 CATCGATGGGAATCTCAGACGACGACCTTTATCATATGACATATTCAGCGAAGTTATGG 923
QY 459 TGACGAGGACCCCGCCACGTACGTGTGGGACGCAACTTACCACGACGCTGGGAAACCGCG 518
Db 924 CAGGACAATCGGCCACGTACATCTGGAAGAACAACTTTTCATGAAAG----- 972
QY 519 GCGCGCGCGCGCGTGGCGCGCGGAGCTATCGGAGGTGGAGCGGAGGATCTCTGGGCG 578
Db 973 -----GGACCTGTGACACAGAGAGTCCCATAAAGAAAGTTC 1010
QY 579 GCTGCGAGTGGAGCGCGTGTGTGAGGACTCGCAGGTGTGTGCGTGGCGCGAGCCTGGGG 638
Db 1011 ACAACATGTACCCCAATTGGTCTTTGTCTAAACAAAGTATTTTGCATCATGCAAAAATGGGG 1070
QY 639 CGCAAGCGGAGGTGATGACACAGCTTCAAGCGCTCGGAGATTCGGTGGACTACGT 698
Db 1071 TGAAGAGAGAGAGTCAATGATACAGCTTTCAGAGCACTGGGGCGAAGCGTGGATTATGT 1130
QY 699 GCAGGTCTGTGACTCGGACCAAGGTTGGACCCCATGGCACTGTGTGAGGTCTGTGCGGGT 758
Db 1131 ACAGGTGTGTGACTCAGATCTATGCTTTGACCTGCTCTCTCTGTGGAGATGGTGAAGGT 1190
QY 759 ACTGAGCAGGACCCCGGAGTGGGCTGTGTGGGAGCGTGGGAGTCTTAAACCTCT 818
Db 1191 CTTAGAGGAGACCCCTATGGTGTGGAGGTGTGGAGGAGATGTCCAGATTTTAAACAAAGTA 1250
QY 819 GGAATCTCTGGGTGAGTCTTCTTAAGCAGCTGCGGATCTGGGTAGCTTCAATGTGGAGCG 878
```


; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15470
; LENGTH: 9871
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15470

Query Match 15.7%; Score 167.6; DB 4; Length 9871;
Best Local Similarity 58.9%; Pred. No. 1.5e-24;
Matches 384; Conservative 0; Mismatches 22; Indels 39; Gaps 4;
QY 61 CGAGGGTGTGACCATCGCCTTCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
DB 2028 CGTGTGTGGGACACGAGCTGTTGCCCTGGCAGTGTGGTGGCATCTGGCAGCCTAT 2087
QY 121 GCCCGGGGTGCGCTGGCTCCGATCGCTAGCGGCTCTGGGCTTGGGCTCTACGGG 180
DB 2088 GTGACGGGCTACCAAGTTTCATCCACACGGAAGCACTACCTGTCTTGGGCTGTACGGC 2147
QY 181 GCCTTCTTTTTCAGCGCACTGTGTGGCGAGAGCCTCTTCCGCTACTGGAGCACCAGCGG 240
DB 2148 GCCATCTTGGGCTGACCTGCTGATTCAGAGCCTTTTTCCTTCTGGAGCACCAGCGGC 2207
QY 241 GTGGCGGGCGGGCGGGCGGCTGTGATGAGCCACCGCGC---GCAGTGTGGCGCTG 297
DB 2208 ATGGAGCTGCGGGCGAGCCCTGAAGCTGCCCTCCCGCGGGGCTCGGTGGCACTG 2267
QY 298 ACCATCTCGGCTTACAGAGAGACCCCGGCTACTGTGGCGAGTGTGGGCTCGGCTCGGCGC 357
DB 2268 TGCATTTGGCGATACAGAGAGACCTGACTACTTGGCGAAGTGTGCTCGGCTCGGCGCAG 2327
QY 358 GCCTGTGTGTACCGCGCGCGGCTGCGCTCTCATGTTGGTGGTGGTGGTGGTGGTGGTGG 417
DB 2328 CGCATCTCTTCCCTGAC-----CTCAAGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2381
QY 418 GAGGACCTTACATGCTGACATGTTCCGCGAGGCTTTCGCTGAGAGAGACCCCGCCACG 477
DB 2382 GAGGAGCCTACATGCTGACATCTTCCAGAGTGTGGGCGGACCGAGCAGCGCGC 2441
QY 478 TACGTGTGGAGCGGAATACACACAGCCCTGGGAAACCCCGCGCGGGCGGCGGCTGGGCG 537
DB 2442 TCTTTGTGTGGCGAGCACTTCAT-----GAGGAGCGGAGGCTGAG 2486
QY 538 GCGGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGGCGCTGGCAGTGGAGCGCTG 597
DB 2487 ACGGAGCGCAGCTTCAGAGGGCATGACCGT-----GTGCGGATGTG 2531
QY 598 GTGAGGACTCGCAGTGTGTGCTGGCGAGCGCTGGGCGGCAAGCGGAGGTCTATG 657
DB 2532 GTGCGGGCCAGACCTTCTCGTGCATCATGACAGAGTGGGAGGCAAGCGGAGGTCTATG 2591
QY 658 TACACAGCCTTCAAGCGCTCGAGATTCGGTGGACTACGTCAGGCTGTG 709
DB 2592 TACACGGCTTCAAGGCCCTCGGCGATTCTGGTGGACTACATCCAGGTAAGG 2643

RESULT 13
US-09-949-016-132199
; Sequence 132199, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132199
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-132199

Query Match 12.9%; Score 137.4; DB 4; Length 601;
Best Local Similarity 58.5%; Pred. No. 1.2e-18;
Matches 300; Conservative 0; Mismatches 201; Indels 12; Gaps 3;
QY 61 CGAGGGTGTGACCATCGCCTTGGCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
DB 50 CGTGTGTGGGACACGAGCTGTTGCCCTGGCAGTGTGGTGGCATCTGGCAGCCTAT 109
QY 121 GCCCGGGGTGCGCTGGCTCCGATCGCTACGCGCTCTTGGGCTTCTGGGCTCTTACGGG 180
DB 110 GTGACGGGCTACCAAGTTTCATCCACACGGAAGCACTACCTGTCTTGGGCTGTACGGC 169
QY 181 GCCTTCTTTTTCAGCGCACTGTGTGGCGAGAGCCTTCTTCCGCTACTGGAGCACCAGCGG 240
DB 170 GCCATCTTGGGCTGCACTGCTCATTCAGAGCCTTTTTCCTTCTGGAGCACCAGCGC 229
QY 241 GTGGCGGGCGGGCGGGCGGCTGTGATGAGCCACCGCGC---GCAGTGTGGCGCTG 297
DB 230 ATGGAGCTGCGGGCAGGCCCTGAAGTGCCTTCCCGCGGGGCTCGGTGGCACTG 289
QY 298 ACCATCTCGGCTTACAGAGAGACCCCGCTACTGTGGCGAGTGTGGGCTCGGCTCGGCGC 357
DB 290 TGCATTTGGCGCTACAGAGAGACCTGACTACTTGGCGAAGTGTGCTCGGCTCGGCGCAG 349
QY 358 GCCTGTGTGTACCGCGCGGGCTGCGCTCTCATGTTGGTGGTGGTGGTGGTGGTGGTGG 417
DB 350 CGCATCTCTTCCCTGAC-----CTCAAGTGTGTGATGGTGGTGGTGGTGGTGGTGGTGG 403
QY 418 GAGGACCTTACATGCTGACATGTTCCGCGAGGCTT---CGCTGAGAGAGACCCCGCC 474
DB 404 GAGAGCCTTACATGCTGGACATCTTCCAGAGTGTGGGCGGACCGAGCAGCGCGC 463
QY 475 ACGTACGTGTGGGACGGCACTACACAGCCCTGGGAAACCCCGCGGGCGGCGGCTGGG 534
DB 464 TCTTTGTGTGGCGAGCACTTCATGAGGAGCGAGGCTGAGACGAGGCGGCGGCGCTG 523
QY 535 GCGCGCGGAGCTATCGGAGGTGGAGCGGAG 567
DB 524 CAGGAGGCGCATGGACCGTGTGCGGGATGTGGT 556

RESULT 14
US-09-949-016-12479
; Sequence 12479, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12479
; LENGTH: 32176
; TYPE: DNA

```
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(32176)
; OTHER INFORMATION: n = A,T,C or G
US-08-675-499A-25

Query Match      12.0%; Score 127.8; DB 4; Length 32176;
Best Local Similarity 67.4%; Pred. No. 1.6e-16;
Matches 180; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 798 CGTGGGATCTTAACCCCTCTGGACTCCTGGGTGAGCTTCTTAAGCAGCCCTGCGATCTG 857
Db 27685 CTTGAGATTTTAAACAAGTACGATTCTGGATCTCATCTCTCAGCAGTGTAAAGATTTG 27744

Qy 858 GGTAGCCTTCAATGTGGAGCGGGCTGTGCAGAGCTACTTCCACTGTGTATCCTGCATCAG 917
Db 27745 GATGGCTTTTAATATAGAAAGGGCTGTCACTCTTATTTTGGGTGTGTTTCAGTGCATTAG 27804

Qy 918 CGGTCTCTAGCCCTATATAGGATTAACCTCTTGCAGCAGTCTTCTGAGGCTGTGTACAA 977
Db 27805 TGGACCTCTGGGAATGTACAGAACTCCTTGTTCATGAGTTTGTGGAAGATTGGTACAA 27864

Qy 978 CCAGAAGTTCCTGGGTACCCACTGTACTTCTTGGGGATGACGGGCACCTCAACCAACCGCAT 1037
Db 27865 TCAAGAATTTATGGCAACCAATGTAGCTTTGGTGATGACAGGCATCTCACGACCGGGT 27924

Qy 1038 GCTCAGCATGGTTATGTACCAAGTA 1064
Db 27925 GCTGAGCCTGGGCTATGCAACAAATA 27951

RESULT 15
US-08-675-499A-25
; Sequence 25, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-675-499A-25

Query Match      9.0%; Score 95.6; DB 4; Length 235;
Best Local Similarity 71.8%; Pred. No. 2.4e-10;
Matches 125; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 891 CTACTTCCACTGTGTATCTCTGATCATCAGCGTCTCTTAGGCCTATATAGGAATAACCTCTT 950
Db 4 CTACTTTGGCTGTGTGCAGTGTATTAGTGGGCCCTTGGGCATGTACCGCAACAGCCTCCT 63

Qy 951 GCACAGTCTTCTGAGGCTGGTACAAACCAAGTCTCTGGGTACCCACTGTACTTTGG 1010
Db 64 CCAGCAGTCTCTGGAGGACTGGTACCATCATCAAGTCTCTTAGGCAGCAAGTGCAGCTTCGG 123

Qy 1011 GGATGACCGGCACCTCAACCAACCGCATGTCTGAGCATGGGTTATGCTACCAAGTA 1064
Db 124 GGATGACCGGCACCTCAACCAACCGAGTCTCTGAGCCTTGGCTACCGAACTAAGTA 177

Search completed: March 13, 2005, 08:12:22
Job time : 212.778 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 08:01:42 ; Search time 749.676 Seconds
(without alignments)
8450.560 Million cell updates/sec

Title: US-10-672-399-7

Perfect score: 1065

Sequence: 1 atgagacagcagcgccgcccc.....tgggtatgtaccacagtaa 1065

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1065	100.0	1065	18	US-10-672-399-7
2	1064	99.9	1737	18	US-10-672-399-1
3	1064	99.9	2116	13	US-10-042-523-1
4	928	87.1	1083	18	US-10-672-399-5
5	927.4	87.1	1071	18	US-10-672-399-3
6	786.6	73.9	1752	10	US-09-902-939-1
7	422	39.6	662	18	US-10-363-345A-13975
8	422	39.6	662	18	US-10-363-345A-13976
9	421	39.5	490	10	US-09-918-995-27210
c	393.8	37.0	662	18	US-10-363-345A-13973
11	393.8	37.0	662	18	US-10-363-345A-13974

12	353	33.1	1662	17	US-10-295-027-369	Sequence 369, App
13	353	33.1	1662	17	US-10-188-832-136	Sequence 136, App
14	323.8	30.4	1767	17	US-10-309-560-9	Sequence 9, Appli
15	311.6	29.3	1665	10	US-09-902-939-3	Sequence 3, Appli
16	246.8	23.2	1659	10	US-09-902-939-2	Sequence 2, Appli
17	246.8	23.2	4194	15	US-10-262-526-3	Sequence 3, Appli
18	221	20.8	3003	10	US-09-918-624B-64	Sequence 64, Appli
19	221	20.8	3003	15	US-10-262-526-1	Sequence 1, Appli
20	221	20.8	4018	10	US-09-814-353-19942	Sequence 19942, A
21	212.4	19.9	3387	15	US-10-084-817-184	Sequence 184, App
22	212.4	19.9	3387	15	US-10-252-157-84	Sequence 84, Appli
23	100.2	9.4	793	9	US-09-910-943-552	Sequence 552, App
24	95.6	9.0	570	9	US-09-880-107-3449	Sequence 3449, Ap
25	72	6.8	15559	18	US-10-646-664-1	Sequence 1, Appli
26	71	6.7	2182	18	US-10-437-963-62269	Sequence 62269, A
c	70	6.6	43058	9	US-09-954-456-529	Sequence 292, App
c	70	6.6	43058	9	US-09-954-456-529	Sequence 529, App
c	70	6.6	43058	9	US-09-880-107-3950	Sequence 3950, Ap
30	69.4	6.5	5784	14	US-10-152-886-64	Sequence 64, Appli
31	68.6	6.4	1755	18	US-10-437-963-78301	Sequence 78301, A
32	68.4	6.4	2214	15	US-10-156-761-2510	Sequence 2510, Ap
c	68.4	6.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
34	67.6	6.3	1380	18	US-10-425-115-170699	Sequence 170699,
35	66.6	6.3	3957	16	US-10-200-562-193	Sequence 193, App
36	66.6	6.3	3957	16	US-10-237-551-193	Sequence 193, App
37	66.6	6.3	154746	10	US-09-827-688-8	Sequence 8, Appli
c	66.6	6.3	154746	10	US-09-827-688-8	Sequence 8, Appli
c	66.2	6.2	1614	9	US-09-976-740-45	Sequence 45, Appli
c	66.2	6.2	1614	13	US-10-023-529-45	Sequence 45, Appli
c	66.2	6.2	1614	13	US-10-023-523-45	Sequence 45, Appli
c	66.2	6.2	1614	17	US-10-616-187-45	Sequence 45, Appli
c	66.2	6.2	1614	17	US-10-671-242-45	Sequence 45, Appli
c	66.2	6.2	12425	9	US-09-976-740-50	Sequence 50, Appli
c	66.2	6.2	12425	13	US-10-023-529-50	Sequence 50, Appli

ALIGNMENTS

RESULT 1

US-10-672-399-7
; Sequence 7, Application US/10672399
; Publication No. US2005000368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-7

Query Match	100.0%	Score	1065;	DB	18;	Length	1065;
Best Local Similarity	100.0%	Pred. No.	2.8e-271;				
Matches	1065;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	ATGAGACAGCAGGCGCGCCCAAGCCCACTCTCGAGCCCGCCGCTGCTCGGCGCTGGCC	60				
Db	1	ATGAGACAGCAGGCGCGCCCAAGCCCACTCTCGAGCCCGCCGCTGCTCGGCGCTGGCC	60				
Qy	61	CGGAGGGTGTGACCATCGCCTTGGCCCTGTCTTCCTCGGCGCTCATGACCTGGCGCTTAC	120				
Db	61	CGGAGGGTGTGACCATCGCCTTGGCCCTGTCTTCCTCGGCGCTCATGACCTGGCGCTTAC	120				
Qy	121	CCGCGCGGGGTGCGCGCTTCCGATCGCTACGCGCTTCTCGGCGCTTCTCGGCGCTTACGGG	180				
Db	121	CCGCGCGGGGTGCGCGCTTCCGATCGCTACGCGCTTCTCGGCGCTTCTCGGCGCTTACGGG	180				

Db 121 GCCCGCGGGTCCGCTGCGCTCCGATCGTACGCGCTCTCGGCCTCTACGGG 180
Qy 181 GCCTTCCTTTAGCGCACTGTGTGGCGCAGAGCCTCTTGGCGTACTGTGAGCACCAGCGG 240
Db 181 GCCTTCCTTTAGCGCACTGTGTGGCGCAGAGCCTCTTGGCGTACTGTGAGCACCAGCGG 240
Qy 241 GTGGCGGGCGGCGCGGGCGCGTGTGATGACGACACCGCGCGAGTGTGCGCTGACC 300
Db 241 GTGGCGGGCGGCGCGGGCGCGTGTGATGACGACACCGCGCGAGTGTGCGCTGACC 300
Qy 301 ATCTCCGCTACACAGAGAACCCCGCTACCTGCGCCAGTGTGCTGCGCGCGCC 360
Db 301 ATCTCCGCTACACAGAGAACCCCGCTACCTGCGCCAGTGTGCTGCGCGCGCC 360
Qy 361 CTGCTGTATACCGCGCGCGCGTGTGCTGCTCATGTGTGTGGATGGCAACCGCGCGGAG 420
Db 361 CTGCTGTATACCGCGCGCGCGTGTGCTGCTCATGTGTGTGGATGGCAACCGCGCGGAG 420
Qy 421 GACCTCTATCATGTGTGCGAGATGTTCCGCGAGGTCTTTCGCTGACGAGAACCCCGCACGTAC 480
Db 421 GACCTCTATCATGTGTGCGAGATGTTCCGCGAGGTCTTTCGCTGACGAGAACCCCGCACGTAC 480
Qy 481 GTGTGGGACGGCAACTACACAGCCCTGGGAACCCCGCGCGCGCGCGTGGCGGCC 540
Db 481 GTGTGGGACGGCAACTACACAGCCCTGGGAACCCCGCGCGCGCGCGTGGCGGCC 540
Qy 541 GGAGCCTATCGGGAGGTGAGCGGAGGATCTTGGCGCGCTGGCGAGTGGAGCGCTGTGTG 600
Db 541 GGAGCCTATCGGGAGGTGAGCGGAGGATCTTGGCGCGCTGGCGAGTGGAGCGCTGTGTG 600
Qy 601 AGGACTCGCAGGTGTGTGCTGCGCAGCGCTGGCGCGCGCGCGCGCGAGGTCTATGTAC 660
Db 601 AGGACTCGCAGGTGTGTGCTGCGCAGCGCTGGCGCGCGCGCGCGAGGTCTATGTAC 660
Qy 661 ACAGCCTTTCAAGCGCCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACA 720
Db 661 ACAGCCTTTCAAGCGCCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACA 720
Qy 721 AGGTTGAGCCCATGGCATCTGTGTGAGCTCTGTGCGGGTACTGGACGAGAACCCCGGGTA 780
Db 721 AGGTTGAGCCCATGGCATCTGTGTGAGCTCTGTGCGGGTACTGGACGAGAACCCCGGGTA 780
Qy 781 GGGGCTGTGTGGGAGCTGCGGATCTTAAACCTCTGGACTCTGGGTGAGTCTTCTTA 840
Db 781 GGGGCTGTGTGGGAGCTGCGGATCTTAAACCTCTGGACTCTGGGTGAGTCTTCTTA 840
Qy 841 AGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGTGTGTGAGTCTGGACACA 900
Db 841 AGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGTGTGTGAGTCTTCCAC 900
Qy 901 TGTGTATCTGATCAGCGGTCTCTAGGCTATATAGGAATAACCTCTTGGAGAGTTT 960
Db 901 TGTGTATCTGATCAGCGGTCTCTAGGCTATATAGGAATAACCTCTTGGAGAGTTT 960
Qy 961 CTTGAGGCTGTGTACAAACAGAGTTCTGGGTACCCACTGACTTTTGGGGATGACCGG 1020
Db 961 CTTGAGGCTGTGTACAAACAGAGTTCTGGGTACCCACTGACTTTTGGGGATGACCGG 1020
Qy 1021 CACCTCACCAACCGCATCTCAGCATGGGTATGTACCAAGTAA 1065
Db 1021 CACCTCACCAACCGCATCTCAGCATGGGTATGTACCAAGTAA 1065

RESULT 2

US-10-672-399-1
; Sequence 1, Application US/10672399
; Publication No. US2005003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25

; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-1

Query Match 99.9%; Score 1064; DB 18; Length 1737;
Best Local Similarity 100.0%; Pred. No. 5.3e-271;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGACGACGAGGAGCGGCCCAAGCCCACTCTCTGAGGCCCGCCGCTGCTCGGCCCTTGSCC 60
Db 1 ATGAGACGACGAGGAGCGGCCCAAGCCCACTCTCTGAGGCCCGCCGCTGCTCGGCCCTTGSCC 60
Qy 61 CGGAGGGTGTGACCATTCGCCCTTCGCCCTGCTCATCTCTGGGCCCTCATGACCTTGGSCCTTAC 120
Db 61 CGGAGGGTGTGACCATTCGCCCTTCGCCCTGCTCATCTCTGGGCCCTCATGACCTTGGSCCTTAC 120
Qy 121 GCCCGCGGGGTGCGCGCTCGGATCGCTACCGGCTCTCTGGCCTTCTCGGCTCTTACGGG 180
Db 121 GCCCGCGGGGTGCGCGCTCGGATCGCTACCGGCTCTCTGGCCTTCTCGGCTCTTACGGG 180
Qy 181 GCCTTCCTTTTTCAGCGCACCTGTGTGGCGCAGAGCCTCTTTCGGGTACTTGGAGCACCAGCGG 240
Db 181 GCCTTCCTTTTTCAGCGCACCTGTGTGGCGCAGAGCCTCTTTCGGGTACTTGGAGCACCAGCGG 240
Qy 241 GTGGCGGGCGGCGCGGGGGCGCTGTGATGACGCCAACCCGCGCGCAGTGTGGCGCTGACC 300
Db 241 GTGGCGGGCGGCGCGGGGGCGCTGTGATGACGCCAACCCGCGCGCAGTGTGGCGCTGACC 300
Qy 301 ATCTCCGCTTACACGAGGAGAACCCCGCGCTACCTGGCGCCAGTGCCTGGCGCTCCGCCGCGCC 360
Db 301 ATCTCCGCTTACACGAGGAGAACCCCGCGCTACCTGGCGCCAGTGCCTGGCGCTCCGCCGCGCC 360
Qy 361 CTGCTGTATACCGCGCGCGCGCTGCGGTCTCTCATGTGTGTGGATGGCAACCGCGCGGAG 420
Db 361 CTGCTGTATACCGCGCGCGCGCTGCGGTCTCTCATGTGTGTGGATGGCAACCGCGCGGAG 420
Qy 421 GACCTCTACATGTGTGCAATGTTCCGCGAGGTCTTTCGCTGACGAGAACCCCGCACAGTAC 480
Db 421 GACCTCTACATGTGTGCAATGTTCCGCGAGGTCTTTCGCTGACGAGAACCCCGCACAGTAC 480
Qy 481 GTGTGGGACGGCAACTACACAGCCCTTGGAAACCCCGCGCGCGCGCGTGGCGGCC 540
Db 481 GTGTGGGACGGCAACTACACAGCCCTTGGAAACCCCGCGCGCGCGCGTGGCGGCC 540
Qy 541 GGAGCCTATCGGGAGGTGAGCGGAGGATCTCTGGCGGGTGGCAGTGGAGGCGCTGTGTG 600
Db 541 GGAGCCTATCGGGAGGTGAGCGGAGGATCTCTGGCGGGTGGCAGTGGAGGCGCTGTGTG 600
Qy 601 AGGACTCGCAGGTGTGTGCTGCGCAGCGCTGGCGCGCGCAAGCGCGAGGTCTATGTAC 660
Db 601 AGGACTCGCAGGTGTGTGCTGCGCAGCGCTGGCGCGCGCAAGCGCGAGGTCTATGTAC 660
Qy 661 ACAGCCTTTCAAGCGCCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACA 720
Db 661 ACAGCCTTTCAAGCGCCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACA 720
Qy 721 AGGTTGAGCCCATGGCATCTGTGTGAGCTCTGTGCGGGTACTGGACGAGAACCCCGGGTA 780
Db 721 AGGTTGAGCCCATGGCATCTGTGTGAGCTCTGTGCGGGTACTGGACGAGAACCCCGGGTA 780
Qy 781 GGGGCTGTGTGGGAGCTGCGGATCTTAAACCTCTGGACTCTGGGTGAGTCTTCTTA 840
Db 781 GGGGCTGTGTGGGAGCTGCGGATCTTAAACCTCTGGACTCTGGGTGAGTCTTCTTA 840
Qy 841 AGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGTGTGTGAGTCTTCCAC 900
Db 841 AGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGTGTGTGAGTCTTCCAC 900

QY 901 TGTGTATCTGCATCAGCGGTCCTCTAGCCCTATATAGGAATAACCTCTTGCAGAGTTT 960
Db 901 TGTGTATCTGCATCAGCGGTCCTCTAGCCCTATATAGGAATAACCTCTTGCAGAGTTT 960
QY 961 CTTGAGGCTGTGTACAAACAGAGTTCTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
Db 961 CTTGAGGCTGTGTACAAACAGAGTTCTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1064
Db 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1064

RESULT 3

US-10-042-523-1
; Sequence 1, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-042-523-1

Query Match 99.9%; Score 1064; DB 13; Length 2116;
Best Local Similarity 100.0%; Pred. No. 5.3e-271; Indels 0; Gaps 0;
Matches 1064; Conservative 0; Mismatches 0;
QY 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCCGCGCTGTCTCGGCTGGCC 60
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCCGCGCTGTCTCGGCTGGCC 95
QY 61 CGGAGGCTGTGACCATCGCCTTCGCTTCATCTCGGCTCATGACCTGGGCTTAC 120
Db 96 CGGAGGCTGTGACCATCGCCTTCGCTTCATCTCGGCTCATGACCTGGGCTTAC 155

QY 121 GCCGCCGGGGTGCCTGCTGGCTCCGATCGCTACGGCCTCTTGGCCTTGGCCTCTACGGG 180
Db 156 GCCGCCGGGGTGCCTGCTGGCTCCGATCGCTACGGCCTCTTGGCCTTGGCCTCTACGGG 215
QY 191 GCCTTCCTTTTTCAGCGCACCTGGTGGCGAGAGCCTCTTTCGGGTACTCTGAGAGACCGGGG 240
Db 216 GCCTTCCTTTTTCAGCGCACCTGGTGGCGAGAGCCTCTTTCGGGTACTCTGAGAGACCGGGG 275
QY 241 GTGGCGGGCGGGCGGGGGCGCTGGATGCAGCACCGCGCGCAGTGTGGCGCTGACG 300
Db 276 GTGGCGGGCGGGCGGGGGCGCTGGATGCAGCACCGCGCGCAGTGTGGCGCTGACG 335
QY 301 ATCTCCGCTTACCAGGAGGACCCCGCTACCTTGGGCCAGTGCCTGGCGCTCCGCCCGCGCC 360
Db 336 ATCTCCGCTTACCAGGAGGACCCCGCTACCTTGGGCCAGTGCCTGGCGCTCCGCCCGCGCC 395
QY 361 CTGCTGTATACCGCGGGCGGGCTGGCGCTCTCATGTGTGGTGGATGGCAACCGCGCCGAG 420
Db 396 CTGCTGTATACCGCGGGCGGGCTGGCGCTCTCATGTGTGGTGGATGGCAACCGCGCCGAG 455
QY 421 GACCTCTACATGCTGCACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACGTAC 480
Db 456 GACCTCTACATGCTGCACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACGTAC 515
QY 481 GTGTGGGACGGCAACTACACAGCCCTTGGGAACCCCGCGGGCGGGCGCGGTGGCGGCC 540
Db 516 GTGTGGGACGGCAACTACACAGCCCTTGGGAACCCCGCGGGCGGGCGCGGTGGCGGCC 575
QY 541 GGAGCCTATCGGAGGTGGAGGGGAGGATCTTGGGGCGGCTGGCAGTGGAGGCGCTGTG 600
Db 576 GGAGCCTATCGGAGGTGGAGGGGAGGATCTTGGGGCGGCTGGCAGTGGAGGCGCTGTG 635
QY 601 AGGACTCCAGGTGCTGTGCGCGCAGCGCTGGGGCGGCAAGCGGAGGTCTATGTAC 660
Db 636 AGGACTCCAGGTGCTGTGCGCGCAGCGCTGGGGCGGCAAGCGGAGGTCTATGTAC 695
QY 661 ACAGCCTTCAAGGGCGCTCGGAGATTTCGGTGGAGTACGTGACGAGGTCTGTGACCTCGGACACA 720
Db 696 ACAGCCTTCAAGGGCGCTCGGAGATTTCGGTGGAGTACGTGACGAGGTCTGTGACCTCGGACACA 755
QY 721 AGTTTGGACCCCATGGGCACTGCTGGAGCTCTGGGGGTACTGGAGAGAGACCCCGGGTA 780
Db 756 AGTTTGGACCCCATGGGCACTGCTGGAGCTCTGGGGGTACTGGAGAGAGACCCCGGGTA 815
QY 781 GGGGCTGTGTGGGAGCTGCGGATCCTTAACCTCTGGACTCCTGGGTCTGAGCTTCTCTA 840
Db 816 GGGGCTGTGTGGGAGCTGCGGATCCTTAACCTCTGGACTCCTGGGTCTGAGCTTCTCTA 875
QY 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCAC 900
Db 876 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCAC 935
QY 901 TGTGTATCTGCATCAGCGGCTCTTAGGCTATATAGGAATAACCTCTTTCAGAGAGTTT 960
Db 936 TGTGTATCTGCATCAGCGGCTCTTAGGCTATATAGGAATAACCTCTTTCAGAGAGTTT 995
QY 961 CTTGAGGCTGTGTACAAACAGAGTTCTCTGGGTACCCACTGACTTTTGGGGATGACCGG 1020
Db 996 CTTGAGGCTGTGTACAAACAGAGTTCTCTGGGTACCCACTGACTTTTGGGGATGACCGG 1055
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1064
Db 1056 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1099

RESULT 4

US-10-672-399-5
; Sequence 5, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics


```
QY 601 AGGACTCGAGTGGTGTGCTGCGCAGCGCTGGGGCGGCAAGCGAGGTCTATGTAC 660
Db 601 AGGACTCGAGTGGTGTGCTGCGCAGCGCTGGGGCGGCAAGCGAGGTCTATGTAC 660
QY 661 ACAGCTTCAAGGCGCTCGGAGATTGCGTGGACTAGTGCAGGTCTGTGACTCGACACA 720
Db 661 ACAGCTTCAAGGCGCTCGGAGATTGCGTGGACTAGTGCAGGTCTGTGACTCGACACA 720
QY 721 AGGTTGGACCCCATGGCACTGCTGAGCTCGTGGGCTACTGGACGAGACCCCGGGTA 780
Db 721 AGGTTGGACCCCATGGCACTGCTGAGCTCGTGGGCTACTGGACGAGACCCCGGGTA 780
QY 781 GGGGCTGTGTGGGGAAGTGGGAGTCTTAACCTCTGGAATCTCTGGGTCACTTCTTA 840
Db 781 GGGGCTGTGTGGGGAAGTGGGAGTCTTAACCTCTGGAATCTCTGGGTCACTTCTTA 840
QY 841 AGCAGCTGCGATACCTGGTACCTCAATGTGGAGCGGCTTGTGAGCTACTTCCAC 900
Db 841 AGCAGCTGCGATACCTGGTACCTCAATGTGGAGCGGCTTGTGAGCTACTTCCAC 900
QY 901 TGTGTATCTGCATCAGCGGCTCTCTAGG 929
Db 901 TGTGTATCTGCATCAGCGGCTCTCTAGG 929
```

RESULT 6

```
US-902-939-1
; Sequence 1, Application US/0902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: mouse
US-902-939-1
```

```
Query Match 73.9%; Score 786.6; DB 10; Length 1752;
Best Local Similarity 84.9%; Pred. No. 8.7e-198;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;
```

```
QY 10 CAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCGGCTTGGCCGCGAGGGTG 69
Db 7 CAGGACATGCCAAAGCCCTCAGAGGCGCGGCTTCTGCTCTGGCTTGGCCAGCGGAGCA 66
QY 70 CTGACCATCGCTTGGCGCTGCTATCTTGGGCTCATGCTGGGCTCATGCTGGGCTATGCGCGGG 129
Db 67 CTCAGATCATCTTTGGCGCTGCTATCTTGGGCTCATGCTGGGCTCATGCTGGGCTATGCGCGAGGC 126
QY 130 GTGCGCTGGCTCGGATCGCTTACGCGCTCTTGGGCTTGGGCTTACGCGGCTTCTCTT 189
Db 127 GTTCTCTGGCTTACGATCGCTATGAGCTTCTGGGCTTGGGCTTATGGGCTATGCGGCTTCTC 186
QY 190 TCAGCGCACCTGGTGGCGAGAGCTCTTTCGCTACCTGGAGCACCGGCGGCTGGCGCG 249
Db 187 AGCGCACACCTAGTGGCACAGAGCTCTTTCGCTTACCTGGAGCACCGGAGGCTGGCAGCG 246
QY 250 GCGGCGCG-----GGGCGGCTTGGATGAGCGCACCGCGCGGCTGAGTGGCG 294
Db 247 GCTGCGCGCGCTCTCTTGGCGAAGGGGCGCTTGGATGAGCGCACCTGCGAGCGGTGCA 306
QY 295 CTGACCATCTCGGCTTACGAGAGGACCGCGTACCTGCGGCTGCGGCTGCGGCTCGGCC 354
Db 307 CTCACCATCTCAGGCTTACGAGAGTCCCGCTTACCTGCGGCTGCGGCTGCGGCTCGGCC 366
QY 355 CGCGCCCTGTGTATCCCGCGCGCGGCTGGCGGCTCTCTCATGTTGGTGGATGGCAACCGC 414
```

```
Db 367 CGCGCTTGTGTATCCCGCACACAGAGTTACGCGTGTCTCATGGTGTGGACGCAACCGC 426
QY 415 GCCGAGGACCTTATACATGTTGACATGTTCCGCGAGGTCTTTCGCTGACGAGAGACCCCGCC 474
Db 427 GCTGAGGATCTGTATCATGTTGGAGCATGTTCCGAGAAGTCTTTCGCGCATGAGAGACCCCGCC 486
QY 475 ACGTACGTGTGGGACGCGCAATACCAACAGCCCTCGGAACCCCGCGCGG---CGGGCGCG 531
Db 487 ACTTATGTTGGATGGCAACTACCATCAGCCCTGGGAACCAACAGCGGAGGTACTCGGGCGCT 546
QY 532 GTGGCGCGCGAGGACCTATCGGGAGGTGAGGCGGAGGATCTCTGGCGGCTGCGAGTGGAG 591
Db 547 GTGCGTGAAGTGGCTTACCGGGAGGTGAGGCGGAGGACCCCGCGGCTTGGCGGTGGAG 606
QY 592 GCGCTGGTGAAGACTCGCAGGTGCGTGGCGCAGCGCTGGCGCGGCGGCGGCGCGAG 651
Db 607 GCGCTGGTGAAGAACACCGCAGGTGCGTGGCGTGGCTCAGCGTTGGGGCGGCAACGTTGAG 666
QY 652 GTCATGTACACAGCCCTTCAAGCGCTCGGAGATTGCGTGGACTACGTGCGAGGTCTGTGAC 711
Db 667 GTCATGTACACAGCTTTCAGGCACTGGGCGACTCCGTTGGACTACGTGCGAGGTCTGTGAC 726
QY 712 TCGGACACAAGGTTGGACCCCATGSCACTGCTGAGACTGCTGCGGCTACTGGAACGAGAC 771
Db 727 TCAGACACAAGACTAGACCCCATGGCACTGCTGGAGCTTGTGCGAGTGTGGATGAAGAC 786
QY 772 CCGCGGTAGGGCTGTTGGTGGGACGTGGGATCTTAAACCTCTCGACTCTCTGGGTC 831
Db 787 CCGCGGTAGGGCTGTTGGAGGGGATGTGAGGATCTTAAACCTCTCGACTCTCTGGGTC 846
QY 832 AGCTTCTTAAGCAGCGCTCGGATACCTGCGTGGTGGCTTCAATGTGGAGCGGCTTGTGAGAC 891
Db 847 AGCTTCTTGAAGCTCTTTCGATCTGCGTGGTGGCTTCAATGTGGAGCGGCTTGTGAGAC 906
QY 892 TACTTCCACTGTGTATCTGTCATCAGCGGTCTCTAGGCTTATATAGGAATAACCTCTTGG 951
Db 907 TACTTCCACTGTGTCTCTGTCATCAGTGGTCTCTGCGGTCTATACAGAAACAATCTCTCTG 966
QY 952 CAGCAGTCTTCTGAGGCTGCTGACCAAGTCTCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1011
Db 967 CAGCAGTCTTCTGAGGCTGCTGACCAAGTCTGTTAGCATGGGCTATGTACCAAGTA 1079
QY 1012 GATGACGCGCACCTCACCACCGCATGCTGAGCATGGGCTTATGTACCAAGTA 1064
Db 1027 GATGACGCGCACCTCACCACCGCATGCTTAGCATGGGCTATGTACCAAGTA 1079
```

RESULT 7

```
US-10-363-345A-13975
; Sequence 13975, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: EOI/1227
; CURRENT APPLICATION NUMBER: US/10/363.345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 13975
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 13975
US-10-363-345A-13975
```

```
Query Match 39.6%; Score 422; DB 18; Length 662;
Best Local Similarity 77.3%; Pred. No. 1.4e-101;
```

Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Qy 34 GCAGCCCGCGCTGCTCCGGCTGCGCCGCGAGGCTGCTGACCATCGCTTCGCCCTGCTC 93
Db |||||
1 GTAGTTTTCGCTGTTTCGGTTTGGTTTGGTTCGGAGGGTGTGATATATCGTTTTCGTTTGT 60
Qy 94 ATCTGGGCGCTCATGACCTGGGCTTACCGCCGCGGGGTGCGCTGCGCTCCGATCGCTAC 153
Db |||||
61 ATTTTGGGTTTATGATTTGGGTTTACGTCGTCGGGGTGTGTTGGTTTTCGATCGTTAC 120
Qy 154 GGCCTCTCGGCTTCGGCTCTACCGGGCTTCCTTTTCAGCGCACCTCGTGGCGCAGAGC 213
Db |||||
121 GGTTTTTTGGTTTTCGGTTTTCACGGGTTTTCGGGTTTTCGGGTTTTCAGCGTATTTGGTGGCGTAGAGT 180
Qy 214 CTCCTCGGCTACCTGGAGCACCGGCGGGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 273
Db |||||
181 TTTTTCGCGTATTTGGAGTATCGGCGGGTGGCGCGCGCGCGCGGGGTGCTGTTGGATGTA 240
Qy 274 GCCACCGCGCGAGTGTGGCTGACCATCTCCGCTTACCAGGAGGACCCCGCGTACCTG 333
Db |||||
241 GTTATCGCGGCTAGTGTGGCTGTGATTTATTTTCGTTTATTTAGGAGATTTTCGCTATTG 300
Qy 334 CGCCAGTCCCTGGGCTCCGCGCGCCCTGCTGTACCCGCGCGCGCTGCGGCTCCTC 393
Db |||||
301 CGTTAGTGTTCGGCTTCGTTTCGGCTTGTGTTGTTATTCGCGCGCGGTTGCGCGTTT 360
Qy 394 ATGCTGTGGATGGCAACCGCGCGGAGGACCTTACATGCTGATGCTTCGCGAGGTC 453
Db |||||
361 ATGCTGTGGATGGTAATCGCGTCCAGGATTTTATATGCTGATGTTTTCGCGAGGTT 420
Qy 454 TTCGCTGACGAGGACCCCGCACGCTGCTGTGGAGCGCAACTACACAGCCCTGGGAA 513
Db |||||
421 TTCGTTGACGAGATTTCTGTTACGTACGTGTGGAGCGTATTTATTTAGTTTGGAA 480
Qy 514 CCCGCGCGCGGGCGCGGTGGCGCGCGGCTGATCGGAGGCTTACGCGGAGGAGCGCGAGATCCT 573
Db |||||
481 TTCGCGCGCGCGCGCGGTGGCGCTCGAGTTTATCGGAGGTGGAGCGCGAGATTTT 540
Qy 574 GGGCGGCTGGCAGTGGAGCGCTGTGTAGGACTCGCAGGTGCGTGGCGCGAGCGC 633
Db |||||
541 GGGCGGTTGGTAGTGGAGCGCTGTGTAGGATTCGTAGTGTGCTGCGTGGCGTAGCGT 600
Qy 634 TGGGCGGCAAGCGCGAGGTATGTACACAGCTTCAAGGGCTCGGAGATTCGCTGGAC 693
Db |||||
601 TGGGCGGTAGCGGAGGTTATGTATATAGTTTAAAGGCTTCGGAGATTCGCTGGAT 660
Qy 694 TA 695
Db |||
661 TA 662

RESULT 8
US-10-363-345A-13976/c
; Sequence 13976, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 13976
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 13976
US-10-363-345A-13976

Query Match 39,6%; Score 422; DB 18; Length 662;
Best Local Similarity 77,3%; Pred. No. 1.4e-101;
Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Qy 34 GCAGCCCGCGCTGCTCCGGCTGCGCCGCGAGGCTGCTGACCATCGCTTCGCCCTGCTC 93
Db 662 GTAGTTTTCGCTGTTTCGGTTTGGTTCGGAGGGTGTGATATATCGTTTTCGTTTGT 603
Qy 94 ATCTGGGCGCTCATGACCTGGGCTTACCGCCGCGGGGTGCGCTGCGCTCCGATCGCTAC 153
Db 602 ATTTTGGGTTTATGATTTGGGTTTACGTCGTCGGGGTGTGCTGTTGGTTTTCGATCGTTAC 543
Qy 154 GGCCTCTCGGCTTCGGCTCTACGGGCTTCCTTTTCAGCGCACCTCGTGGCGCAGAGC 213
Db 542 GGTTTTTTGGTTTTCGGTTTTCAGGGGTTTTTTTTTTAGCGTATTTGGTGGCGTAGAGT 483
Qy 214 CTCCTCGGCTACCTGGAGCACCGCGCGGTGCGCGCGCGGGGGGGGGGGGGGGGGGGGG 273
Db 482 TTTTTCGCGTATTTGGAGTATCGCGGGTGGCGCGCGCGGGGTGCTGTTGGATGTA 423
Qy 274 GCCACCGCGCGAGTGTGGGCTGACCATCTCCGCTTACAGGAGGACCCCGCGTACCTG 333
Db 422 GTTATCGCGCGTAGTGTGGCGTTGATTTATTTTCGTTTATTTAGGAGGATTTTCGCGTATTG 363
Qy 334 CGCCAGTCCCTGGGCTCCGCGCGCCCTGCTGTACCCGCGCGCGCTGCGGCTCCTC 393
Db 362 CGTTAGTGTTCGGGCTGCTGTTTCGGCTTTGTTGTTATTCGCGCGCGGTTGCGCGTTTT 303
Qy 394 ATGCTGTGGATGGCAACCGCGCGGAGGACCTTACATGCTGATGCTTCGCGAGGTC 453
Db 302 ATGCTGTGGATGGTAATCGCGTCCAGGATTTTATATGCTGATATGTTTCGCGAGGTT 243
Qy 454 TTCGCTGACGAGGACCCCGCACGCTGCTGTGGAGCGCAACTACACAGCCCTGGGAA 513
Db 242 TTCGTTGACGAGGATTTTCGTTACGTACGTGTGGAGCGTAAATTTATTTAGTTTGGGAA 183
Qy 514 CCCGCGCGCGCGCGGTGGCGCGCGGCTTACGCGAGGCTTACGCGAGGAGGAGGATCCT 573
Db 182 TTCGCGCGCGCGCGCGGTGGCGCTCGAGTTTATCGGAGGTGGAGCGCGAGGATTTT 123
Qy 574 GGGCGGCTGGCAGTGGAGCGCTGCTGAGGACTCGCAGGTCGCTGCTGCGCGCAGCGC 633
Db 122 GGGCGGTTGGTAGTGGAGCGCTGCTGAGGATTCGTAGTGTGCTGCGTGGCGTAGCGT 63
Qy 634 TGGGCGGCAAGCGCGAGGTCATGTACACAGCTTCAAGGGCTCGGAGATTCGCTGGAC 693
Db 62 TGGGCGGTAAAGCGGAGGTTATGTATATAGTTTTTAAAGGCTTCGGAGATTCGCTGGAT 3
Qy 694 TA 695
Db |||
2 TA 1

RESULT 9
US-09-918-995-27210
; Sequence 27210, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27210

Query Match 39.5%; Score 421; DB 10; Length 490;
Best Local Similarity 90.9%; Pred. No. 2.6e-101;
Matches 442; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 544 GCCTATCGGAGGTGCGAGCGGAGATCTGGCGGCTGGCAGTGGAGCGCGTGGTGGG 603
DB 2 CGAGTAGGGAAANNNNNNNAATTGAGACTGAGGCACNGTCGGGAANTCGCTGGTGGG 61
QY 604 ACTCGCAGGTGCTGTGCTGCGCAGCGCTGGGCGCGCAAGCGGAGTCTATGTACACA 663
DB 62 ACTCGCAGGTGCTGTGCTGCGCAGCGCTGGGCGCGCAAGCGGAGTCTATGTACACA 121
QY 664 GCCTTCAAGCGCTCGGAGATTGCTGTGACGTACGTGCAAGGTCTGTGACTCGGACACAAGG 723
DB 122 GCCTTCAAGCGCTCGGAGATTGCTGTGACGTACGTGCAAGGTCTGTGACTCGGACACAAGG 181
QY 724 TTGGACCCCATGGCACTCTGGAGCTCTGCGGCTGCTGGGCTGCTGGAGCGGACCCCGGTAGG 783
DB 182 TTGGACCCCATGGCACTCTGGAGCTCTGCGGCTGCTGGGCTGCTGGAGCGGACCCCGGTAGG 241
QY 784 GCTGTGTGGGGAGCTCGGATCCTTAACTCTGGACTCTGGGCTGAGTCTTCTTAAGC 843
DB 242 GCTGTGTGGGGAGTGTGGGATCCTTAACTCTGGACTCTGGGCTGAGTCTTCTTAAGC 301
QY 844 AGCTTGCATACTGGGTAGCTTCAATGTGAGCGGGTGTGTCAGAGTACTTCCACTGT 903
DB 302 AGCTTGCATACTGGGTAGCTTCAATGTGAGCGGGTGTGTCAGAGTACTTCCACTGT 361
QY 904 GTATCTTCATCAGCGGCTCTAGGCTTATAGGCTTATAGGATTAACCTTTCGACGAGTTTCT 963
DB 362 GTATCTTCATCAGCGGCTCTAGGCTTATAGGCTTATAGGATTAACCTTTCGACGAGTTTCT 421
QY 964 GAGGCTTGTACACAGAGTTCTGGGTACCCACTGCTACTTTTGGGGATGACCGGCAC 1023
DB 422 GAGGCTTGTACACAGAGTTCTGGGTACCCACTGCTACTTTTGGGGATGACCGGCCT 481
QY 1024 CTCACC 1029
DB 482 CACAAC 487

RESULT 10
US-10-363-345A-13973/c
; Sequence 13973, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 13973
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-Island No: 13973
US-10-363-345A-13973

Query Match 37.0%; Score 393.8; DB 18; Length 662;
Best Local Similarity 74.7%; Pred. No. 4e-94;
Matches 494; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

```

```

QY 35 CAGCCCGCGCTGTCTCGGCTTGGCCCGAGGGTGTGACCATCGCCTTGCCTCTGCTCA 94
DB 661 CAACCTACCGCTACTCCGACCTAACCCGAAATAATAACCATCGCTTGCCTTACTCA 602
QY 95 TCCTGGGCTCATGACCTGGGCTTACGCGCGGGGTGCGTGGCTTCCGATCGCTAGG 154
DB 601 TCCTAAACCTTCAATACCTTAAACCTTACGCGCGGGAATACCGCTAACCTCCGATCGCTAGG 542
QY 155 GCCTCTGGCTTGGGCTTACGCGGCTTCTCTTTTACGCGCACCTGTGTGGCGAGAGCC 214
DB 541 ACCTCTAACCTTTCGACCTTACGAAACCTTCTCTTTTCAACGCACTTAATAACGCAAAACC 482
QY 215 TCTTGGCTTACTTGAGCACCGGCGGGTGGGCGCGCGCGCGCGCGCTGTGGATGCGAG 274
DB 481 TCTTGGCTTACTTAAACACACCGAGAAATAACGACGACGCGGAAACCGCTTAATACAA 422
QY 275 CCACGCGCGGAGTGTGGGCTGACCATCTCGGCTTACAGGAGGACCCCGGTACTGTC 334
DB 421 CCACGCGCGCAATATATACGCTTAAACCTTCTCGGCTTACAAACCCCGGTACTTAC 362
QY 335 GCCAGTGTCTGCGCTCGGCGCGCGCTTGTGTACCGCGCGCGCGCTTCCGCGCTCTCA 394
DB 361 GCCAATCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 302
QY 395 TGGTGTGGATGGCAACCGCGCGCGCTTCTACATGCTGACATGTTTCCGCGAGGTCT 454
DB 301 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 242
QY 455 TCGCTGACGAGGACCCCGCCACGTCAGTGTGGGAGCGCACTACACAGCCCTGGGAAC 514
DB 241 TCGCTTAAAGAAACCCCGCCACGTCAGTGTATATAACGCACTTACCAACCCCTTAAAC 182
QY 515 CCGCGCGCGCGCGCGCTGGGCGCGGAGCTTACCGGAGTGTGGAGGAGGATCCCTG 574
DB 181 CCGGACGACGAGCGGATTAACCGCGGAACTTATCGAAATAATAATAATAATAATAATA 122
QY 575 GCGCGCTGGAGTGGAGCGCTGGTGGAGTCTCGAGGTGGTGTGCTGGCGGAGCGCT 634
DB 121 AACGACTTAACTTAAACCGCTTAAATAAACTCGCAATACGTATACGTAAAGCAACGCT 62
QY 635 GGGCGGCAAGCGGAGTGTATGTACAGGCTTCAAGCGGCTCGGAGATTCGGTGGACT 694
DB 61 AAAACGACAAACGCGAATCATATACACACCTTCAAAACGCTCGAAATTCGATAAACT 2
QY 695 A 695
DB 1 A 1

```

```

RESULT 11
US-10-363-345A-13974
; Sequence 13974, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 13974
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-Island No: 13974
US-10-363-345A-13974

```



```
Query Match      37.0%; Score 393.8; DB 18; Length 662;
Best Local Similarity 74.7%; Pred. No. 4e-94;
Matches 494; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

Qy 35 CAGCCCGCGCTGCTCCGGCTTGGCCGGAGGCTGCTGACCATGCGCTTGGCTGCTGCTCA 94
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2 CAACCTACCGCTACTCCGACCTTAACCCGAAATAATAACCATGCGCTTGGCCCTACTCA 61
Qy 95 TCCTGGGCTCATGACCTGGGCTACGCGCGGGGTCGCGCTGCGCTCCGATCGCTACG 154
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 TCCTAAACCTATAACCTAAACCTACGCGCGGAAATACCGCTAACCTCCGATCGCTACG 121
Qy 155 GCCTCTCGGCTTCCGGCTCTACGGGGCTTCTCTTTCAGCGCACCTGGTGGCGAGGCC 214
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 ACCTCTAACCTTGAACCTTACGAAACCTTCTTTCACGCACTTATAAGCGAAAC 181
Qy 215 TCTTCGCTACTGAGACCGCGGGTGGCGGGCGCGGGCGGGCGCTGGATGAG 274
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 TCTTCGCTACTAAACACCGACGAATAACGACGACGCGAAACCGCTAAATACAA 241
Qy 275 CCACCGCGGAGTGGCGGTGACCATCTCGCTTACCGCGCGCGCTGCGCTGCTCA 334
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
242 CCACCGCGCAATATAACGCTAACCATCTCGCTTACCAAAACCCCGCTACCTAC 301
Qy 335 GCCAGTCTGCGCTCCCGCGCGCTGCTGTATCCCGCGCGCGCTGCGCTGCTCA 394
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
302 GCCAATACCTAACGCTCCCGCGCGCTTACTATACCGCGCGCGCTACGCTCTCA 361
Qy 395 TGGTGGTGGATGGCAACCGCGCGAGGACCTTCTACATGGTGCATGTTCCGCGAGTCT 454
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 TAATAATAATAACCAACCGCGCGGAAACCTTCTACATAATGACATATTCGCGGAAATCT 421
Qy 455 TCGTGTGAGGAGCCCGCCACGTACGTGGGAGGCACTACACACGCGCTGGGAAC 514
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
422 TCGTGTACGAAACCCCGCCACGTACGTATAAAACGACAACTACCAACCCCTAAAAAC 481
Qy 515 CCGCGCGCGCGCGCGCTGGCGCGCGGAGCTATCGGAGGTGGAGCGGAGATCTGT 574
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
482 CCGGACGACGACGAGTAACCGCGGAAACCTATCGAAATAATAAACGAAATCTTA 541
Qy 575 GCGGCTGGAGTGGAGCGCTGTTGAGGACTGCGAGGTGGTGGTGGCGGAGCGCT 634
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
542 AACGACTAACCAATAAAACCGCTTAATAAAACCTCGCAATAGCTATACGTAACGCAACGCT 601
Qy 635 GGGCGGCAACGCGAGGTCTATGACAGGCTTCAAGCGCTCGGAGATTCGTTGACT 694
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
602 AAAACGACAAACGCGAAATCATATACAAACCTTCAAAACGCTCGAAATTCGATAAAT 661
Qy 695 A 695
Db 662 A 662
```

RESULT 12
US-10-295-027-369
; Sequence 369, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13

```
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 369
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-369
```

Query Match 33.1%; Score 353; DB 17; Length 1662;
Best Local Similarity 62.9%; Pred. No. 2.4e-83;
Matches 633; Conservative 0; Mismatches 335; Indels 39; Gaps 4;

```
Qy 61 CGGAGGCTGTGACCATCGCTTCCGCTTGTCTATCTGGGCTCATGACCTGGGCTTAC 120
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
28 CGTGTGTGGCACCAACGCTTGTGGCTTGGGCTGCGATCTCTGGAGCTCTCTGGAGCTAT 87
Qy 121 GCGCGCGGGTGGCGCTCGGCTCGGCTACGCGCTTCCGCTTCCGCTTCTACGGG 180
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 GTGACGGGCTACCAAGTTCATCCACGGAANAACACTACTGTCTTCTGGCTGTACGGC 147
Qy 181 GCCTTCTCTTTCAGCGCACTGGTGGCGAGAGCTTCTTTCGCTTACCTGGAGACCGGCGG 240
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
148 GCCATCTGGGCTGACCTGCTCATTCAGAGCTTTTGGCTTCTTGGAGCACCGGCGC 207
Qy 241 GTGCGGCGGGCGCGCGGGGCGCTGGATGACACCAACCGCGC---GCAGTGTGGCGCTG 297
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
208 ATGCGAGCTGCGCGCCAGGCGCTGAAGCTGCCCTCCCGCGCGGGGCTCGGTGGCACTG 267
Qy 298 ACCATCTCCGCTTACCAAGGAGGACCCCGCGTACCTGGCGCAAGTGCCTGGCGCTCGCGCGC 357
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
268 TGCATTGCGCATACCAAGGAGGACCTGACTACTTGGCAAGTGCCTGGCTCGGCTCGGCCAG 327
Qy 358 GCCCTGTGTACCGCGCGCGGCTGGCTCTCATGTGTGGTGGATGGCAACCGCGCC 417
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
328 CGCATCTCTTCCCTGAC-----CTCAAGGTGTGTGTGGTGGATGGCAACCGGCCAG 381
Qy 418 GAGGACCTTACATGTTGCGAGAGTCTTCCGCGAGGTCTTCTGCTGACGAGGACCCCGCCACG 477
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
382 GAGGACGCTTACATGCTGGGACATCTTCAAGAGTGTGGGCGGACCGAGAGCGCGCCG 441
Qy 478 TACGTGTGGAGCGGCAACTACCAACAGCCCTGGGAAACCCCGCGGGCGGGCGCGGTGGGC 537
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
442 TTCTTTGTGGCGCAGCAACTTCCA-----TGAGGACAGGAGGGGTGAG 486
Qy 538 GCGGAGCCTTATCGGAGGTGGAGCGGAGGATCCTGGGCGGCTGGCAGTGGAGGCGCTG 597
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
487 ACGGAGCGACCTGACGAGGCGCATGGACCGT-----GTGCGGATGTG 531
Qy 598 GTGAGGACTCGAGGTGGTGTGGTGGCGCGCTGGGCGGCAAGCGCGAGGTGATG 657
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
532 GTGCGGGCCAGCACCTTCTGTGTCATCATGCAAGATGGGGAGGCAAGCGCGAGGTGATG 591
```



```

; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Xenopus laevis
US-10-309-560-9

```

Query Match	30.4%	Score 323.8	DB 17	Length 1767
Best Local Similarity	58.5%	Pred. No. 1.2e-75		
Matches	587	Conservative 0	Mismatches 407	Indels 9
	Gaps	1		
Qy	62	GGAGGGTGTGACCATCGCCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCGCTACG	121	
Db	80	GGAGGATAATTTATTTCTTTTGGTGGTGTATTAGCTTACCATTACAGACGCTATG	139	
Qy	122	CCGCGGGGTGCCGCTCGCCTCCCATCGCTACGGCCCTCTGGCCCTCTGGCCCTCTACGGGG	181	
Db	140	TGGCAGAGTCCAGGTCTCAAAATGAAGCCATCTCTCTCTCTGGGCTTTATGGTC	199	
Qy	182	CTTCTCTTTTCAGCGCACTGGTGGCGAGAGCCTCTTCGCGTATCTGTGAGCACCGCGGG	241	
Db	200	TTGCAATGCTTCTCCACTGATGATGAGAGCCTCTTTGCTTCTCTGGAGATACGCGGG	259	
Qy	242	TGGCGGGCGGGCGGGGGCGCTGGATGTCAGCACCGCGCGCAGTGTGCGCTGACCA	301	
Db	260	TAAATAAGATGA-----GCTTCTTGCAGCTTTAAGAACAGTAGTGGCTCTGACCA	310	
Qy	302	TCTCCGCTTACCAGGAGGACCCCGCTACCTTCGCCAGTGGCTGGCGTCCGCCCGCGCC	361	
Db	311	TTGCTGGGTATCAGAGAACCTTGAGTACCTGATATAAGTGTCTGGAATCTGTCAAGTATG	370	
Qy	362	TGCTGTATCCCGCGCGCGGGTGGCGTCTCTCATGTGTGGTGGATGGCAACCGCGCCGAGG	421	
Db	371	TGAAATACCCCAAGATAAACTCAAGATCATTTTGGTTCATCGATGGGAAACACAGAGGATG	430	
Qy	422	ACCTCTACATGTTGACATGTTCCGCGAGTCTTCGCTGACGAGGACCCCGCCACGTAACG	481	
Db	431	ATGCCTACATGATGGAGATGTTTCAAGACGTGTTCCAGGTGAAGATGTAGGACCTTACG	490	
Qy	482	TGTGGGACGGCAACTACACAGCCCTTGGGAACCCGCGGGCGGGCGGTGGGCGCGG	541	
Db	491	TATGGAGGGAATTACCACACTGTTTAAAGCCTGAGGAGACCAATAAGGATCCTGTCT	550	
Qy	542	GAGCCTATCCGGAGGTGGAGCGAGGATCTCTGGCGGCTGGCAGTGGAGGCGCTGTGTGA	601	
Db	551	CTGAGGTTTCTAAGCCCTTGAAATGAAGATGAAGGTATCAATATGTTGGGAAGAACTTGTTA	610	
Qy	602	GGACTCGCAGTGTGCTGGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCTATGTACA	661	
Db	611	GAAACAAGAGATGTGTGTGATCATGTAACAGTGGGGCGGAAAAGAGAGGTCTATGTACA	670	
Qy	662	CAGCCTTTAAGGGGCTCGGAGATTCGGTGGACTACGTGCAGGTCGTGTCACTCGGACACAA	721	
Db	671	CAGCATTCAGGCCATTGGGACTTCTGTGGACTATGTACAGGTCTGTGACTCGGACACCA	730	
Qy	722	GTTTGGACCCCATGGCACTGCTGGAGCTCTGTGGGGTACTGACAGAGACCCCGGGTAG	781	
Db	731	AACTGGATGAATGGCAACAGTGGAAATGGTGAAGGTTCTGGAATCCAATGACATGTATCG	790	
Qy	782	GGGCTGTTGGTGGGAGTGGGATCCTTAACCTCTGGAATCTGGTCTGGGTGAGTTCCTAA	841	
Db	791	GCGCAGTGGGAGAGAGTTCGCAATCTGAACCCCTTATGATTCCTTCATTAGTTTCATGA	850	
Qy	842	GCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGTGTGTACAGAGTACTTCCACT	901	
Db	851	GCAGCCTGCGTTACTGGATGGGTTTAACTGGAGCGGGTGTGTACAGAGTACTTCCACT	910	
Qy	902	GTGTATCTCGATCAGGGGTCTCTAGGCCCTATATAGGAATAACCTCTTGCAGCAGTTTC	961	
Db	911	GCGTGTCTGTATAAGTGGACCTCTGGGAATGTACCGGAACAAATCTCCAGGTGTTTT	970	

Qy	962	TTGAGGCGCTGGTACACAGAAAGTTCTCTGGGTACCCACTGTACTCTTTTGGGGATGACCGGC	1021
Db	971	TGGAAGCCTCGTTCAGACAAAAATTTTGGGAACCTATTGTACTTTGGGAGATGATAGAC	1030
Qy	1022	ACCTCACCAACCCGATGCTCAGCATGGTTATGCTACCAAGTA	1064
Db	1031	ATCTGACAAACCGAGTGCTCAGCATGGATATCGACCAATA	1073
RESULT 15			
US-09-902-939-3			
; Sequence 3, Application US/09902939			
; Publication No. US20030087850A1			
; GENERAL INFORMATION:			
; APPLICANT: Philip DeHaza			
; APPLICANT: William Chen			
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME			
; FILE REFERENCE: 2055/OH20-US0			
; CURRENT APPLICATION NUMBER: US/09/902,939			
; CURRENT FILING DATE: 2001-07-10			
; NUMBER OF SEQ ID NOS: 6			
; SOFTWARE: Fast-Seq for Windows Version 3.0			
; SEQ ID NO 3			
; LENGTH: 1665			
; TYPE: DNA			
; ORGANISM: mouse			
US-09-902-939-3			
Query Match 29.3%; Score 311.6; DB 10; Length 1665;			
Best Local Similarity 60.3%; Pred. No. 2.1e-72;			
Matches 609; Conservative 0; Mismatches 359; Indels 42; Gaps 4;			
Qy	61	CGGAGGCTGTGACCATCGCTTCGGCCCTGCTCATCTCTGGGCGCTCATGACCTCTGGGCGCTAC	120
Db	28	CGTGTGGTGGGACACAGTCTGTTTGGCCCTGGTAGTCTCTGGAGGCACTCTGGGCGGCTAT	87
Qy	121	GCGCCCGGGGTGCCGTGGCTCCGATCGCTACGGCTCTCTGGCCCTTCCTGGCCCTCTACGGG	180
Db	88	GTGACAGGCTACACAGTTTATCCACACAGAAAGCACTACCTGTCTCTTGGCCCTCTACGGT	147
Qy	181	GCCTTCCTTTCAGGCGACCTGGTGGCGCAGAGGCTCTTTCGCGTACTCTGAGNACCCGGCGG	240
Db	148	GCCATCTCTGGGTCTATCATCTGCTCATCCAGAGGCTGTTTGCCTTCTCTGGAGCACCTCGA	207
Qy	241	GTGCGCGCGCGCGCGG-----GGGCGCGCTGGATGCAGCCACCGCGCGCAGTGTGGCG	294
Db	208	ATGCGCAGGCGAGGGCGCCCTCAAGCTGCATCTGCCAGAGGTCTGGGCTTCACTGGCA	267
Qy	295	CTGACCATCTCCGCTTACAGGAGGACCCCGGTACTCTGCGCAGTGCCTCTGGCGTCCGCC	354
Db	268	CTCTGCATTGCTGCCTTACCAGAGGAGCCCGAATACCTGCGCAAGTGCCTTCGCTCAGCT	327
Qy	355	CGGCGCTGTGTATCCGCGCGCGCGGCTGCGGCTCTCATGTGTGTGATGTCACACCGC	414
Db	328	CAGCGCATTTGCCCTTCCAAAC-----CTCAAGGTGTGTATGGTAGTGTGATGTCATCGC	381
Qy	415	GCCGAGGACCTTACATGTTCGACATGTTCCGCGAGGTCTTCGCTCACGAGGACCCCGCC	474
Db	382	CAGAAAGATACCTTACATGTTGGACATCTTCCATGAGGTGCTTGGGTGGCACTGAGCAAGCT	441
Qy	475	ACGTAGCTGTGGGACGGCACTACCAACGACCCCTGGGAACCCCGCGCGCGGCGCGGTG	534
Db	442	GGCTTCTTTGTGTGGCGTAGCAATTTCCA-----TGAGCGGGTGAAGGA	486
Qy	535	GGCGCGGAGCCCTATCGGAGGTGTGAGGCGGAGGATCTCTGGCGGCTGGCAGTGTGAGGCG	594
Db	487	GAGACAGAGCCAGCCTGCGAGGAAGCATGGAGCGT-----GTGCGAGCT	531
Qy	595	CTGGTGAGGACTCGCAGGTGCTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCAGGTC	654
Db	532	GTGGTGTGGGCGCAGCACCTTCTCATGCATCATGCAGAGTGTGGGGGGGCAAGCGTAGGTC	591
Qy	655	ATGTACACAGCCTTCAAGCGGCTTCGAGAGATTCCGTGGCACTACGTGCAAGTCTGTGACTCG	714

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:11:12 ; Search time 5176.05 Seconds
(without alignments)
8399.558 Million cell updates/sec

Title: US-10-672-399-7
Perfect score: 1065
Sequence: 1 atgagacagcaggacgcgc.....tgggtatgtaccacagtaa 1065

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues
Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/PCTUS3_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/PCTUS4_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/PCTUS5_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/PCTUS6_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/PCTUS7_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/PCTUS8_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/PCTUS9_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/PCTUS10_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/PCTUS11_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/PCTUS12_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/PCTUS13_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/PCTUS14_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/PCTUS15_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/PCTUS16_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/PCTUS17_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/PCTUS18_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/PCTUS19_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/PCTUS20_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/PCTUS21_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/PCTUS22_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/PCTUS23_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/PCTUS24_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/PCTUS25_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/PCTUS26_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/PCTUS27_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/PCTUS28_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/PCTUS29_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/PCTUS30_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/PCTUS31_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/PCTUS32_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/PCTUS33_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/PCTUS34_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/PCTUS35_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/PCTUS36_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/PCTUS37_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/PCTUS38_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/PCTUS39_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/PCTUS40_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/PCTUS41_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/PCTUS42_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/PCTUS43_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US110_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq.*
93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*

117: /cgn2_6/ptodata/1/pna/US6047 COMB.seq.*
118: /cgn2_6/ptodata/1/pna/US6048 COMB.seq.*
119: /cgn2_6/ptodata/1/pna/US6049 COMB.seq.*
120: /cgn2_6/ptodata/1/pna/US6050 COMB.seq.*
121: /cgn2_6/ptodata/1/pna/US6051 COMB.seq.*
122: /cgn2_6/ptodata/1/pna/US6052 COMB.seq.*
123: /cgn2_6/ptodata/1/pna/US6053 COMB.seq.*
124: /cgn2_6/ptodata/1/pna/US6054 COMB.seq.*
125: /cgn2_6/ptodata/1/pna/US6055 COMB.seq.*
126: /cgn2_6/ptodata/1/pna/US6056 COMB.seq.*
127: /cgn2_6/ptodata/1/pna/US6057 COMB.seq.*
128: /cgn2_6/ptodata/1/pna/US6058 COMB.seq.*
129: /cgn2_6/ptodata/1/pna/US6059 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1065	100.0	1065	58	US-10-672-399-7	Sequence 7, Appli
2	1064	99.9	1737	1	PCT-US02-41225A-223	Sequence 223, App
3	1064	98.9	1737	58	US-10-672-399-1	Sequence 1, Appli
4	1064	99.9	1737	95	US-60-257-537-223	Sequence 223, App
5	1064	99.9	2088	1	PCT-US02-41225A-222	Sequence 222, App
6	1064	99.9	2088	52	US-10-342-164-1	Sequence 1, Appli
7	1064	99.9	2088	95	US-60-257-537-222	Sequence 222, App
8	1064	99.9	2098	97	US-60-278-258-16271	Sequence 16271, A
9	1064	99.9	2116	46	US-10-042-523-1	Sequence 1, Appli
10	1062.4	99.8	1737	1	PCT-US02-41225A-224	Sequence 224, App
11	1062.4	99.8	1737	95	US-60-257-537-224	Sequence 224, App
12	1062.4	99.8	2087	1	PCT-US02-41225A-226	Sequence 226, App
13	1062.4	99.8	2087	95	US-60-257-537-226	Sequence 226, App
14	1060.8	99.6	2087	49	US-10-170-235-37298	Sequence 37298, A
15	1060.8	99.6	2087	66	US-10-940-774-5659	Sequence 5659, Ap
16	1060.8	99.6	2087	115	US-60-452-680-12253	Sequence 12253, A
17	1060.8	99.6	2087	115	US-60-453-050-7646	Sequence 7646, Ap
18	1060.8	99.6	2087	115	US-60-453-135-7646	Sequence 7646, Ap
19	1060.8	99.6	2087	116	US-60-466-412-7646	Sequence 7646, Ap
20	1060.8	99.6	2089	67	US-10-930-328-7459	Sequence 7459, Ap
21	1060.8	99.6	2099	67	US-10-930-328-7460	Sequence 7460, Ap
22	1011.6	95.0	2112	102	US-60-324-185-23093	Sequence 23093, A
23	928	87.1	1071	58	US-10-672-399-5	Sequence 5, Appli
24	927.4	87.1	1071	58	US-10-672-399-3	Sequence 3, Appli
25	788.2	74.0	2099	48	US-10-144-771-12143	Sequence 12143, A
26	788.2	74.0	2099	106	US-60-360-207-12143	Sequence 12143, A
27	786.6	73.9	1752	1	PCT-US01-21785-1	Sequence 1, Appli
28	786.6	73.9	1752	39	US-09-902-939-1	Sequence 1, Appli
29	697.2	65.5	14862	66	US-10-940-774-17401	Sequence 17401, A
30	697.2	65.5	231222	1	PCT-US02-41225A-225	Sequence 225, App
31	697.2	65.5	231222	95	US-60-257-537-225	Sequence 225, App
32	697.2	65.5	4622116	41	US-09-947-916-288	Sequence 288, App
33	696.4	65.4	22862	116	US-60-466-412-87508	Sequence 87508, A
34	696.4	65.4	22877	67	US-10-990-328-96152	Sequence 96152, A
35	422	39.6	662	53	US-10-363-345A-13975	Sequence 13975, A
36	422	39.6	662	53	US-10-363-345A-13976	Sequence 13976, A
37	422	39.6	662	53	US-10-363-483A-13975	Sequence 13975, A
38	422	39.6	662	53	US-10-363-483A-13976	Sequence 13976, A
39	421	39.5	490	18	US-09-235-076-27210	Sequence 27210, A
40	421	39.5	490	19	US-09-289-768-29320	Sequence 29320, A
41	421	39.5	490	20	US-09-332-782-27210	Sequence 27210, A
42	421	39.5	490	33	US-09-737-223-27210	Sequence 27210, A
43	421	39.5	490	39	US-09-918-995-27210	Sequence 27210, A
44	421	39.5	490	40	US-09-939-397-29320	Sequence 29320, A
45	411.8	38.7	445	23	US-09-528-409-68253	Sequence 68253, A

ALIGNMENTS

RESULT 1
US-10-672-399-7
; Sequence 7, Application US/10672399
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; PRIORITY FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIORITY FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-7

Query Match 100.0%; Score 1065; DB 58; Length 1065;
Best Local Similarity 100.0%; Pred. No. 2e-164;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTGTGAGCCCGCCGCTGTCTCGGCTGTGCCC 60
DB 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTGTGAGCCCGCCGCTGTCTCGGCTGTGCCC 60

QY 61 CGGAGGGGTGTGACATATGCGCTTTCGCTGTCTATCTCTGTGGGCTCATGACCTGGGCTTAC 120
DB 61 CGGAGGGGTGTGACATATGCGCTTTCGCTGTCTATCTCTGTGGGCTCATGACCTGGGCTTAC 120

QY 121 GCCCGCGGGGTGCGCTGGCTTCGATCGCTTACCGGCTTCTGGCTTCTGGGCTTCTACGGG 180
DB 121 GCCCGCGGGGTGCGCTGGCTTCGATCGCTTACCGGCTTCTGGCTTCTGGGCTTCTACGGG 180

QY 181 GCCTTCTCTTTCAGCGCACCTGTGGTGGCGCAGAGCCTTCTCGGCTTACCTGGAGCACCGGGG 240
DB 181 GCCTTCTCTTTCAGCGCACCTGTGGTGGCGCAGAGCCTTCTCGGCTTACCTGGAGCACCGGGG 240

QY 241 GTGCGCGGGCGCGCGCGGGGCGCTGTGATGACGCCACCGCGCGAGGTGTGGGCTGTGACC 300
DB 241 GTGCGCGGGCGCGCGCGGGGCGCTGTGATGACGCCACCGCGCGAGGTGTGGGCTGTGACC 300

QY 301 ATCTCCGCTTACAGGAGGACCGCGCTTACCTGCGGCTTCTGGCTTCTGGGCTTCTACGGG 360
DB 301 ATCTCCGCTTACAGGAGGACCGCGCTTACCTGCGGCTTCTGGCTTCTGGGCTTCTACGGG 360

QY 361 CTGCTGTACCGCGCGCGCGCTGCGGCTTCTCATGTGTGTGGATGGCAACCGCGCGGAG 420
DB 361 CTGCTGTACCGCGCGCGCGCTGCGGCTTCTCATGTGTGTGGATGGCAACCGCGCGGAG 420

QY 421 GACCTTACATGTGTGATGATTTCCGCGAGGTCTTTCGTGTGACGAGGACCCCGCACGTGAC 480
DB 421 GACCTTACATGTGTGATGATTTCCGCGAGGTCTTTCGTGTGACGAGGACCCCGCACGTGAC 480

QY 481 GTGTGGACGCGCACTACCAACCGCTTGGGAAACCGCGGGGCGGGGCGGGTGGGCGCC 540
DB 481 GTGTGGACGCGCACTACCAACCGCTTGGGAAACCGCGGGGCGGGGCGGGTGGGCGCC 540

QY 541 GGAGCTTATCGGAGGTGTGAGCGGAGGATCTCTGGGCGGCTGGGAGGTGGAGGCGCTGTGTG 600
DB 541 GGAGCTTATCGGAGGTGTGAGCGGAGGATCTCTGGGCGGCTGGGAGGTGGAGGCGCTGTGTG 600

QY 601 AGGACTCGAGGTGCGTGTGCGTGGCGAGCGCTTGGGGCGGCAACCGCGAGGTGATGTAC 660
DB 601 AGGACTCGAGGTGCGTGTGCGTGGCGAGCGCTTGGGGCGGCAACCGCGAGGTGATGTAC 660

QY 661 ACAGCTTCAAGGCGCTCGGAGGATTCGGTGTGACTACCTGTGAGTCTGTGACTCGGACACA 720
DB 661 ACAGCTTCAAGGCGCTCGGAGGATTCGGTGTGACTACCTGTGAGTCTGTGACTCGGACACA 720

QY 721 AGGTGTGACCCCATGGCAGCTGTGCGGCTGTGCGGCTGTGCGGAGGACCCCGGGGTA 780
DB 721 AGGTGTGACCCCATGGCAGCTGTGCGGCTGTGCGGCTGTGCGGAGGACCCCGGGGTA 780

301	ATCTCCGCTTACAGAGAGACCCCGGTACTCTGCGCAAGTCTGCGCTCGCGCCGCGCC	360
361	CTGCTGTACCCGCGCGCGGCTGCGCGTCTCTCATGTTGGTGGATGGCAACCGCGCCGAG	420
361	CTGCTGTATCCGCGCGCGGCTGCGCGTCTCTCATGTTGGTGGATGGCAACCGCGCCGAG	420
421	GRACCTCATGTTGCACATGTTCCGCGAGGTTCTGCTGACGAGAGACCCCGCCACGTAC	480
421	GACCTCTCATGTTGCACATGTTCCGCGAGGTTCTGCTGACGAGAGACCCCGCCACGTAC	480
481	GTGTGGGACGGCAACTACCAACGACCCCTGGGAACCCGCGCGCGCGCGTGGCGCC	540
481	GTGTGGGACGGCAACTACCAACGACCCCTGGGAACCCGCGCGCGCGCGTGGCGCC	540
541	GGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGCGACGTGGAGGCGCTGGTG	600
541	GGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGCGACGTGGAGGCGCTGGTG	600
601	AGGACTCGCAGGTGCGTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCACTGTAC	660
601	AGGACTCGCAGGTGCGTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCACTGTAC	660

1111
CACA / 20

661	ACAGCCTTCAAGCGCGCTCGGAGNTTCGGTGGACTACTGTCAGGCTCTGTGACTCGGACACA	720
721	AGGTTGAGACCCCATAGCACTGCTGGAGCTCGTGCGGGTACTGGACGAGACACCCCGGGTA	780
721	AGGTTGAGACCCCATAGCACTGCTGGAGCTCGTGCGGGTACTGGACGAGACACCCCGGGTA	780
781	GGGGCTGTTGGTGGGAGCGTGGGATCCTTAACCCCTCTGGACTCCTGGGTGACGTTCCCTA	840
781	GGGGCTGTTGGTGGGAGCGTGGGATCCTTAACCCCTCTGGACTCCTGGGTGACGTTCCCTA	840
841	AGCAGCGCTCGGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGTCAGAGCTACTTCCAC	900
841	AGCAGCGCTCGGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGTCAGAGCTACTTCCAC	900
901	TGTGTATCTTCGCATCAGCGGTCTCTAGGCCTATATAGGAATAACCTCTTTGCAGCAGTTT	960
901	TGTGTATCTTCGCATCAGCGGTCTCTAGGCCTATATAGGAATAACCTCTTTGCAGCAGTTT	960
961	CTTGAGGCTTGGTACAACAGAAAGTTCTGGGGTACCCACTGTACTTTTGGGGATGACCGG	1020
961	CTTGAGGCTTGGTACAACAGAAAGTTCTGGGGTACCCACTGTACTTTTGGGGATGACCGG	1020

```

0Y      1021 CACCTCACCAACCGCATGCTCAGCATGGTATGTACCAGA 1064
        |||||
Db      1021 CACCTCACCAACCGCATGCTCAGCATGGTATGTACCAGA 1064
        |||||

RESULT 3
US-10-672-399-1
; Sequence 1, Application US/10672399
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-1

Query Match          99.9%; Score 1064; DB 58; Length 1737;
Best Local Similarity 100.0%; Pred. No. 2.8e-164;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCCGCGCTGCTCGGCTGGCC 60
Db 1 ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCCGCGCTGCTCGGCTGGCC 60
QY 61 CGAGGGTGTGACCATCGCTTCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db 61 CGAGGGTGTGACCATCGCTTCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
QY 121 GCCCGCGGGGTCCCGCTCCGATCGCTACCGGCTCTCTGGCTTCCGGCTTACGGG 180
Db 121 GCCCGCGGGGTCCCGCTCCGATCGCTACCGGCTCTCTGGCTTCCGGCTTACGGG 180
QY 181 GCCTTCCTTTAGCGCACCTGGTGGCGAGAGCTTCTGCGGTACCTGGAGCACCGGGG 240
Db 181 GCCTTCCTTTAGCGCACCTGGTGGCGAGAGCTTCTGCGGTACCTGGAGCACCGGGG 240
QY 241 GTGGCGGCGGCGCGGGGCGCTGTGATGAGCCACCGCGCAGTGTGGCGCTGACC 300
Db 241 GTGGCGGCGGCGCGGGGCGCTGTGATGAGCCACCGCGCAGTGTGGCGCTGACC 300
QY 301 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGCCTGGCGTCCGCGGCC 360
Db 301 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGCCTGGCGTCCGCGGCC 360
QY 361 CTGCTGTACCCGCGCGCGGTGGCGTCTCTCATGTGTGGTGGTAAACCGCGCCGAG 420
Db 361 CTGCTGTACCCGCGCGCGGTGGCGTCTCTCATGTGTGGTGGTAAACCGCGCCGAG 420
QY 421 GACCTCTACATGTTGCGATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCACGTAC 480
Db 421 GACCTCTACATGTTGCGATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCACGTAC 480
QY 481 GTGTGGGACGGCAATACCAACCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCC 540
Db 481 GTGTGGGACGGCAATACCAACCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCC 540
QY 541 GGAGCCTATCGGAGGTGGAGCGAGATCTCTGGCGGCTGGAGTGGAGCGCGCTGGTG 600
Db 541 GGAGCCTATCGGAGGTGGAGCGAGATCTCTGGCGGCTGGAGTGGAGCGCGCTGGTG 600
QY 601 AGGACTCGCAGTGTGTGGCGCAGCGCTGGGGCGCGCAAGCGCGAGTCTATGTAC 660
Db 601 AGGACTCGCAGTGTGTGGCGCAGCGCTGGGGCGCGCAAGCGCGAGTCTATGTAC 660
QY 661 ACAGCCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGACTCGGACA 720
Db 661 ACAGCCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGACTCGGACA 720
QY 721 AGGTTGGACCCCATGGCACTGCTGGAGCTCGTGGGCTGCTGGAGAGACCCCGGGTA 780
Db 721 AGGTTGGACCCCATGGCACTGCTGGAGCTCGTGGGCTGCTGGAGAGACCCCGGGTA 780
QY 781 GGGGCTGTGTGGGAGCTGGGATCTTAACCTCTGGAATCTCTGGCTCAGCTTCTTA 840
Db 781 GGGGCTGTGTGGGAGCTGGGATCTTAACCTCTGGAATCTCTGGCTCAGCTTCTTA 840
QY 841 AGCAGCTTCGATATCGGCTAGCTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCAC 900
Db 841 AGCAGCTTCGATATCGGCTAGCTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCAC 900
QY 901 TGTGTATCTGTGATCAGCGGCTCTTAGGCTATATAGGAATAACCTCTTGAGAGGTTT 960
Db 901 TGTGTATCTGTGATCAGCGGCTCTTAGGCTATATAGGAATAACCTCTTGAGAGGTTT 960
QY 961 CTTGAGGCTGTGTACCAAGAGTTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
Db 961 CTTGAGGCTGTGTACCAAGAGTTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
QY 1021 CACCTCACCAACCGCATCTCAGCATGGGTTATGCTACCAAGTA 1064
Db 1021 CACCTCACCAACCGCATCTCAGCATGGGTTATGCTACCAAGTA 1064

RESULT 4
US-60-257-537-223
; Sequence 223, Application US/60257537
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen
; APPLICANT: Schafer, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GX-0022 P
; CURRENT APPLICATION NUMBER: US/60/257,537
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 223
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GB:NM_001523.1
US-60-257-537-223
Query Match 99.9%, Score 1064, DB 95, Length 1737;
Best Local Similarity 100.0%; Pred. No. 2.8e-164;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCCGCGCTGCTCGGCTGGCC 60
Db 1 ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCCGCGCTGCTCGGCTGGCC 60
QY 61 CGAGGGTGTGACCATCGCTTCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db 61 CGAGGGTGTGACCATCGCTTCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
QY 121 GCCCGCGGGGTCCCGCTCCGATCGCTACCGGCTCTCTGGCTTCCGGCTTACGGG 180
Db 121 GCCCGCGGGGTCCCGCTCCGATCGCTACCGGCTCTCTGGCTTCCGGCTTACGGG 180
QY 181 GCCTTCCTTTAGCGCACCTGGTGGCGAGAGCTTCTGCGGTACCTGGAGCACCGGGG 240
Db 181 GCCTTCCTTTAGCGCACCTGGTGGCGAGAGCTTCTGCGGTACCTGGAGCACCGGGG 240
QY 241 GTGCGCGCGCGCGCGGGGCGCTGGATGACGCCACCGCGCGCAGTGTGGCGCTGACC 300
Db 241 GTGCGCGCGCGCGCGGGGCGCTGGATGACGCCACCGCGCGCAGTGTGGCGCTGACC 300
QY 301 ATCTCCGCTTACAGAGAGACCCCGCGTACCTGCGCCAGTGCCTGGGCTCCGCGGCC 360
Db 301 ATCTCCGCTTACAGAGAGACCCCGCGTACCTGCGCCAGTGCCTGGGCTCCGCGGCC 360
QY 361 CTGCTGTACCCGCGCGCGGTGGCGTCTCTCATGTGTGGTGGTAAACCGCGCCGAG 420
Db 361 CTGCTGTACCCGCGCGCGGTGGCGTCTCTCATGTGTGGTGGTAAACCGCGCCGAG 420
QY 421 GACCTCTACATGTTGCGATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCACGTAC 480
Db 421 GACCTCTACATGTTGCGATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCACGTAC 480
QY 481 GTGTGGGACGGCAATACCAACCGCGCGGCGCGCGCGCGCGCGCGCGCGCC 540
Db 481 GTGTGGGACGGCAATACCAACCGCGCGGCGCGCGCGCGCGCGCGCGCGCC 540
QY 541 GGAGCCTATCGGAGGTGGAGCGAGATCTCTGGCGGCTGGAGTGGAGCGCGCTGGTG 600
Db 541 GGAGCCTATCGGAGGTGGAGCGAGATCTCTGGCGGCTGGAGTGGAGCGCGCTGGTG 600
QY 601 AGGACTCGCAGTGTGTGGCGCAGCGCTGGGGCGCGCAAGCGCGAGTCTATGTAC 660
Db 601 AGGACTCGCAGTGTGTGGCGCAGCGCTGGGGCGCGCAAGCGCGAGTCTATGTAC 660
QY 661 ACAGCCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGACTCGGACA 720
Db 661 ACAGCCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGACTCGGACA 720

QY 721 AGTTGGACCCCATGGCACTGCTGGAGCTCGTGGGGTACTGGACGAGAACCCCGGGTA 780
DB 721 AGTTGGACCCCATGGCACTGCTGGAGCTCGTGGGGTACTGGACGAGAACCCCGGGTA 780
QY 781 GGGGTGTTGGTGGGAGCGTGGGATCCTTAACCCCTCTGGACTCTGGGTACAGCTTCCTA 840
DB 781 GGGGTGTTGGTGGGAGCGTGGGATCCTTAACCCCTCTGGACTCTGGGTACAGCTTCCTA 840
QY 841 AGCAGCTGCGATACCTGGGTAGCTTCAATGTGGAGCGGGTGTGTCAGAGCTACTTCCAC 900
DB 841 AGCAGCTGCGATACCTGGGTAGCTTCAATGTGGAGCGGGTGTGTCAGAGCTACTTCCAC 900
QY 901 TGTGTATCTGTCATCAGCGGCTCTCTAGGCCCTATATAGGAATAACCTCTTTCAGCAGTTT 960
DB 901 TGTGTATCTGTCATCAGCGGCTCTCTAGGCCCTATATAGGAATAACCTCTTTCAGCAGTTT 960
QY 961 CTTGAGGCTGGTACAAACAGAGTTCTGGGTACCCACTGACTTTTGGGGATACCGG 1020
DB 961 CTTGAGGCTGGTACAAACAGAGTTCTGGGTACCCACTGACTTTTGGGGATACCGG 1020
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1064
DB 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1064

RESULT 5

PCT-US02-41225A-222
; Sequence 222, Application PC/TUS0241225A
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: JONES, Karen Anne
; APPLICANT: SCHAFER, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GX-0022 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/41225A
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/342,603
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 222
; LENGTH: 2088
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: GB:NM_001523
PCT-US02-41225A-222

Query Match 99.9%; Score 1064; DB 1; Length 2088;
Best Local Similarity 100.0%; Pred. No. 2.8e-164; Mismatches 0; Indels 0; Gaps 0;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACGACGAGCGCCCAAGCCCACTCTGTCAGCGCGCGCTGCTCGGCTCGGCC 60
DB 36 ATGAGACGACGAGCGCGCCCAAGCCCACTCTGTCAGCGCGCGCTGCTCGGCTCGGCC 95
QY 61 CGGAGGTGCTGACCATCGCTTCCCTGCTGCTATCTTGGGCTCATGACCTGGGCTTAC 120
DB 96 CGGAGGTGCTGACCATCGCTTCCCTGCTGCTATCTTGGGCTCATGACCTGGGCTTAC 155
QY 121 GCCCGCGGGTGGCGCTGGCTCCGATCGTACGGCTCTGGGCTCTGGGCTCTACGGG 180
DB 156 GCCCGCGGGTGGCGCTGGCTCCGATCGTACGGCTCTGGGCTCTGGGCTCTACGGG 215
QY 181 GCCTTCCTTTTCAGCGCACCTGGTGGCGCAGAGCTCTTTTCGGTACTTGGAGCACCGCGG 240
DB 216 GCCTTCCTTTTCAGCGCACCTGGTGGCGCAGAGCTCTTTTCGGTACTTGGAGCACCGCGG 275
QY 241 GTGGCGGGCGGGCGCGGGGGCGCTGGATGACGACCAACCGCGCAGTGTGGCGCTGACC 300
DB 276 GTGGCGGGCGGGCGCGGGGGCGCTGGATGACGACCAACCGCGCAGTGTGGCGCTGACC 335

QY 301 ATCTCCGCTACACGAGGAGAACCCCGCTACCTGCGCCAGTGCCTGGCGTCCGCGCCGCGC 360
DB 336 ATCTCCGCTACACGAGGAGAACCCCGCTACCTGCGCCAGTGCCTGGCGTCCGCGCCGCGC 395
QY 361 CTGCTGTATACCCGCGCGCGCGCTGGCGTCTCTCATGTGTGGTGGATGGCAACCGCGCCGAG 420
DB 396 CTGCTGTATACCCGCGCGCGCGCTGGCGTCTCTCATGTGTGGTGGATGGCAACCGCGCCGAG 455
QY 421 GACCTCTACATGGTTCGACATGTTCCGCGAGGCTTTCGCTGACGAGGAGCCCGCCACGTTAC 480
DB 456 GACCTCTACATGGTTCGACATGTTCCGCGAGGCTTTCGCTGACGAGGAGCCCGCCACGTTAC 515
QY 481 GTGTGGGACGGCAACTACACACAGCCCTTGGGAAACCCCGCGCGCGCGCGCTGGCGCC 540
DB 516 GTGTGGGACGGCAACTACACACAGCCCTTGGGAAACCCCGCGCGCGCGCTGGCGCC 575
QY 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCTCTGGCGCGCTGGAGTGGAGGCGCTGGTG 600
DB 576 GGAGCCTATCGGAGGTGGAGCGGAGGATCTCTGGCGCGCTGGAGTGGAGGCGCTGGTG 635
QY 601 AGGACTCGCAGTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCACTGTAC 660
DB 636 AGGACTCGCAGTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCACTGTAC 695
QY 661 ACAGCCTTCAAGCGCGCTCGGAGATTTCGGTGGAGTACGTGCGAGGTCTGTCACTCGGACACA 720
DB 696 ACAGCCTTCAAGCGCGCTCGGAGATTTCGGTGGAGTACGTGCGAGGTCTGTCACTCGGACACA 755
QY 721 AGTTGGACCCCATGGCACTGCTGGAGCTCGTGGGGTACTGGACGAGAACCCCGGGTA 780
DB 756 AGTTGGACCCCATGGCACTGCTGGAGCTCGTGGGGTACTGGACGAGAACCCCGGGTA 815
QY 781 GGGGTGTTGGTGGGAGCGTGGGATCCTTAACCCCTCTGGACTCTGGGTGAGCTTCCTA 840
DB 816 GGGGTGTTGGTGGGAGCGTGGGATCCTTAACCCCTCTGGACTCTGGGTGAGCTTCCTA 875
QY 841 AGCAGCTGCGATACCTGGGTAGCCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCAC 900
DB 876 AGCAGCTGCGATACCTGGGTAGCCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCAC 935
QY 901 TGTGTATCTGTCATCAGCGGCTCTTAGGCCCTATATAGGAATAACCTCTTTCAGCAGTTT 960
DB 936 TGTGTATCTGTCATCAGCGGCTCTTAGGCCCTATATAGGAATAACCTCTTTCAGCAGTTT 995
QY 961 CTTGAGGCTGGTACAAACAGAGTTCTCTGGGTACCCACTGACTTTTGGGGATACCGG 1020
DB 996 CTTGAGGCTGGTACAAACAGAGTTCTCTGGGTACCCACTGACTTTTGGGGATACCGG 1055
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1064
DB 1056 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1099

RESULT 6

US-10-342-164-1
; Sequence 1, Application US/10342164
; GENERAL INFORMATION:
; APPLICANT: McCarthy, James B.
; APPLICANT: Simpson, Melanie A.
; APPLICANT: Bullard, Kelli M.
; APPLICANT: Furcht, Leo T.
; TITLE OF INVENTION: METHODS OF INHIBITING PROSTATE, COLON, BREAST, AND OVARIAN TUMORS
; FILE REFERENCE: 110.0212010
; CURRENT APPLICATION NUMBER: US/10/342,164
; CURRENT FILING DATE: 2003-01-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2088
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-10-342-164-1

Query Match									
Best Local Similarity 99.9%; Score 1064; DB 52; Length 2088;									
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ATGAGACGACGAGGACGCGCCCAAGCCCACTCTCTGACGCCCGCGCTGCTCCGGGCTGGCC	60						
Db	36	ATGAGACGACGAGGACGCGCCCAAGCCCACTCTCTGACGCCCGCGCTGCTCCGGGCTGGCC	95						
Qy	61	CGGAGGGTGCTGACCATCGCCTTCGCCCTGCTCATCTCTGGCCCTCATGACCTGGGCTTAC	120						
Db	96	CGGAGGGTGCTGACCATCGCCTTCGCCCTGCTCATCTCTGGCCCTCATGACCTGGGCTTAC	155						
Qy	121	GCCGCGGGGTGCGCTGCGCTCCGATCGCTACCGGCTCTGGGCTTCCGGCTTACGGG	180						
Db	156	GCCGCGGGGTGCGCTGCGCTCCGATCGCTACCGGCTCTGGGCTTCCGGCTTACGGG	215						
Qy	181	GCTTCTCTTTTCAGCGCACCTGGTGCGCAGAGCCTCTTCGCGTACCTGGAGCACCGGCGG	240						
Db	216	GCTTCTCTTTTCAGCGCACCTGGTGCGCAGAGCCTCTTCGCGTACCTGGAGCACCGGCGG	275						
Qy	241	GTGCG	300						
Db	276	GTGCG	335						
Qy	301	ATCTCCGCTTACAGGAGGACCCCGCGTACCTGCGCCAGTCCCTGGGCTCCGGGCTTACGGG	360						
Db	336	ATCTCCGCTTACAGGAGGACCCCGCGTACCTGCGCCAGTCCCTGGGCTCCGGGCTTACGGG	395						
Qy	361	CTGCTGTATACCG	420						
Db	396	CTGCTGTATACCG	455						
Qy	421	GACCTCTACATGCTGCGATGTTCCGCGAGGCTTTCGCTGACGAGGACCCCGCGCACGTTAC	480						
Db	456	GACCTCTACATGCTGCGATGTTCCGCGAGGCTTTCGCTGACGAGGACCCCGCGCACGTTAC	515						
Qy	481	GTGTGGGACGGCAATACACAGCCCTTGGGAAACCCGCGCGCGCGCGCGCGCGCGCGCGCG	540						
Db	516	GTGTGGGACGGCAATACACAGCCCTTGGGAAACCCGCGCGCGCGCGCGCGCGCGCGCGCG	575						
Qy	541	GGAGCCTATCGGAGGTGGAGCGGAGGATCCTTGGGCGGCTGGCAGTGGAGCGCTGGTG	600						
Db	576	GGAGCCTATCGGAGGTGGAGCGGAGGATCCTTGGGCGGCTGGCAGTGGAGCGCTGGTG	635						
Qy	601	AGGACTCGAGGTGCGTGTGCGTGGCGAGCGCTGGGCGGCGCAAGCGCGAGGTCTATGTAC	660						
Db	636	AGGACTCGAGGTGCGTGTGCGTGGCGAGCGCTGGGCGGCGCAAGCGCGAGGTCTATGTAC	695						

Query Match									
Best Local Similarity 100.0%; Pred. No. 2.8e-164;									
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ATGAGACGACGAGGACGCGCCCAAGCCCACTCTCTGACGCCCGCGCTGCTCCGGGCTGGCC	60						
Db	36	ATGAGACGACGAGGACGCGCCCAAGCCCACTCTCTGACGCCCGCGCTGCTCCGGGCTGGCC	95						
Qy	61	CGGAGGGTGCTGACCATCGCCTTCGCCCTGCTCATCTCTGGCCCTCATGACCTGGGCTTAC	120						
Db	96	CGGAGGGTGCTGACCATCGCCTTCGCCCTGCTCATCTCTGGCCCTCATGACCTGGGCTTAC	155						
Qy	121	GCCGCGGGGTGCGCTGCGCTCCGATCGCTACCGGCTCTGGGCTTCCGGCTTACGGG	180						
Db	156	GCCGCGGGGTGCGCTGCGCTCCGATCGCTACCGGCTCTGGGCTTCCGGCTTACGGG	215						
Qy	181	GCTTCTCTTTTCAGCGCACCTGGTGCGCAGAGCCTCTTCGCGTACCTGGAGCACCGGCGG	240						
Db	216	GCTTCTCTTTTCAGCGCACCTGGTGCGCAGAGCCTCTTCGCGTACCTGGAGCACCGGCGG	275						
Qy	241	GTGCG	300						
Db	276	GTGCG	335						
Qy	301	ATCTCCGCTTACAGGAGGACCCCGCGTACCTGCGCCAGTCCCTGGGCTCCGGGCTTACGGG	360						
Db	336	ATCTCCGCTTACAGGAGGACCCCGCGTACCTGCGCCAGTCCCTGGGCTCCGGGCTTACGGG	395						
Qy	361	CTGCTGTATACCG	420						
Db	396	CTGCTGTATACCG	455						
Qy	421	GACCTCTACATGCTGCGATGTTCCGCGAGGCTTTCGCTGACGAGGACCCCGCGCACGTTAC	480						
Db	456	GACCTCTACATGCTGCGATGTTCCGCGAGGCTTTCGCTGACGAGGACCCCGCGCACGTTAC	515						
Qy	481	GTGTGGGACGGCAATACACAGCCCTTGGGAAACCCGCGCGCGCGCGCGCGCGCGCGCGCG	540						
Db	516	GTGTGGGACGGCAATACACAGCCCTTGGGAAACCCGCGCGCGCGCGCGCGCGCGCGCGCG	575						
Qy	541	GGAGCCTATCGGAGGTGGAGCGGAGGATCCTTGGGCGGCTGGCAGTGGAGCGCTGGTG	600						
Db	576	GGAGCCTATCGGAGGTGGAGCGGAGGATCCTTGGGCGGCTGGCAGTGGAGCGCTGGTG	635						
Qy	601	AGGACTCGAGGTGCGTGTGCGTGGCGAGCGCTGGGCGGCGCAAGCGCGAGGTCTATGTAC	660						
Db	636	AGGACTCGAGGTGCGTGTGCGTGGCGAGCGCTGGGCGGCGCAAGCGCGAGGTCTATGTAC	695						

RESULT 7
US-60-257-537-222
; Sequence 222, Application US/60257537
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen
; APPLICANT: Schafer, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GX-0022 P
; CURRENT APPLICATION NUMBER: US/60/257,537
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 222
; LENGTH: 2088
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GB:NM_001523
US-60-257-537-222

QY 661 ACAGCCTTCAAGCGCTCGAGATTCGGTGGACTAGTGCAGGTCTGTGACTCGGACACA 720
| | | | |
Db 696 ACAGCCTTCAAGCGCTCGAGATTCGGTGGACTAGTGCAGGTCTGTGACTCGGACACA 755
| | | | |
QY 721 AGTTGGACCCCATGCACTGTGGAGCTGTGGGGTACTGGACAGAGACCCCGGGTA 780
| | | | |
Db 756 AGTTGGACCCCATGCACTGTGGAGCTGTGGGGTACTGGACAGAGACCCCGGGTA 815
| | | | |
QY 781 GGGGCTTGTGGTGGGACGTGGGATCTTAACCTCTGAGCTTCCTGGGTTCAGCTTCCTTA 840
| | | | |
Db 816 GGGGCTTGTGGTGGGACGTGGGATCTTAACCTCTGAGCTTCCTGGGTTCAGCTTCCTTA 875
| | | | |
QY 841 AGAGCCTCGGATACCTGGGTAGCTTCATATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
| | | | |
Db 876 AGAGCCTCGGATACCTGGGTAGCTTCATATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 935
| | | | |
QY 901 TGTGTATCTGTCATCAGCGGCTCTAGGCTATATAGAAATACCTCTTGCAGCGATTT 960
| | | | |
Db 936 TGTGTATCTGTCATCAGCGGCTCTAGGCTATATAGAAATACCTCTTGCAGCGATTT 995
| | | | |
QY 961 CTTGAGGCTGTGTACAAACAGAGTTCTCTGGGTACCACTGTACTTTTGGGGATGACCGG 1020
| | | | |
Db 996 CTTGAGGCTGTGTACAAACAGAGTTCTCTGGGTACCACTGTACTTTTGGGGATGACCGG 1055
| | | | |
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGGTATGTGTACCAAGTA 1064
| | | | |
Db 1056 CACCTCACCAACCGCATGCTCAGCATGGGTATGTGTACCAAGTA 1099
| | | | |

RESULT 8

US-60-278-259-16271
; Sequence 16271, Application US/60278258
; GENERAL INFORMATION:
; APPLICANT: Morris, Macdonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; TITLE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0010-1 P
; CURRENT APPLICATION NUMBER: US/60/278,258
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 17730
; SOFTWARE: PERL Program
; SEQ ID NO 16271
; LENGTH: 2098
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 335023.1
US-60-278-259-16271

Query Match 99.9%; Score 1064; DB 97; Length 2098;

Best Local Similarity 100.0%; Pred. No. 2.8e-164;

Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACAGCAGCGCGCCAGCGCCACTCTGTGAGCGCGCGCTGTCTCGGCTGGCC 60
| | | | |
Db 48 ATGAGACAGCAGCGCGCCAGCGCCACTCTGTGAGCGCGCGCTGTCTCGGCTGGCC 107
| | | | |
QY 61 CGAGGGTGTGACCATCGCTTCGCCCTGTCTATCCTGGGCTCATGACCTGGGCGCTAC 120
| | | | |
Db 108 CGAGGGTGTGACCATCGCTTCGCCCTGTCTATCCTGGGCTCATGACCTGGGCGCTAC 167
| | | | |
QY 121 GCCCGCGGGTCCGCTGCGCTCCGATCGCTACGGCTCTTGGCCTTACGGG 180
| | | | |
Db 168 GCCCGCGGGTCCGCTGCGCTCCGATCGCTACGGCTCTTGGCCTTACGGG 227
| | | | |
QY 181 GCCTTCTTTAGCGACCTGTGGGCGAGAGCTCTTGGCGTACTGTGAGCAGCGGG 240
| | | | |
Db 228 GCCTTCTTTAGCGACCTGTGGGCGAGAGCTCTTGGCGTACTGTGAGCAGCGGG 287
| | | | |

RESULT 9

US-10-042-523-1

; Sequence 1, Application US/10042523

; GENERAL INFORMATION:

; APPLICANT: Briskin, Michael J.

; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02173

QY 241 GTGGGGCGCGCGCGGGGCGGTGATGAGCCACCGCGCGCAGTGTGGCGCTGACC 300
| | | | |
Db 288 GTGGGGCGCGCGCGGGGCGGTGATGAGCCACCGCGCGCAGTGTGGCGCTGACC 347
| | | | |
QY 301 ATCTCCGCCCTACAGAGAGACCCCGGCTACTGTGCGCACTGTGCTGCGGCGCGCC 360
| | | | |
Db 348 ATCTCCGCCCTACAGAGAGACCCCGGCTACTGTGCGCACTGTGCTGCGGCGCGCC 407
| | | | |
QY 361 CTGCTGTACCCCGCGCGCGGCTGCGGCTCTCATGTTGGTGGATGCGCAACCGCGCGG 420
| | | | |
Db 408 CTGCTGTACCCCGCGCGCGGCTGCGGCTCTCATGTTGGTGGATGCGCAACCGCGCGG 467
| | | | |
QY 421 GACCTCTACATGGTGCAGCATGTTCCGCGCAGAGTCTTCGTGAGAGAGACCCCGCCAGTAC 480
| | | | |
Db 468 GACCTCTACATGGTGCAGCATGTTCCGCGCAGAGTCTTCGTGAGAGAGACCCCGCCAGTAC 527
| | | | |
QY 481 GTGTGGAGCGCACTACACACAGCCCTGGGAAACCGCGCGCGCGCGCGGCTGGCGGCC 540
| | | | |
Db 528 GTGTGGAGCGCACTACACACAGCCCTGGGAAACCGCGCGCGCGCGCGGCTGGCGGCC 587
| | | | |
QY 541 GGAGCTATCTCGGAGGTGAGCGGAGGATCTCTGGGCGGCTGGCAGTGGAGCGCTGGTG 600
| | | | |
Db 588 GGAGCTATCTCGGAGGTGAGCGGAGGATCTCTGGGCGGCTGGCAGTGGAGCGCTGGTG 647
| | | | |
QY 601 AGGACTCGCAGGTGCTGTGCGCGCAGCGCTCTGGGGCGCAAGCGCGAGGTCTATGTATC 660
| | | | |
Db 648 AGGACTCGCAGGTGCTGTGCGCGCAGCGCTCTGGGGCGCAAGCGCGAGGTCTATGTATC 707
| | | | |
QY 661 ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTAGTGCAGGTCTGTGACTCGGACACA 720
| | | | |
Db 708 ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTAGTGCAGGTCTGTGACTCGGACACA 767
| | | | |
QY 721 AGTTGGACCCCATGCACTGTGAGCTCTGTGCGGGTACTGGACGAGACCCCGGGTA 780
| | | | |
Db 768 AGTTGGACCCCATGCACTGTGAGCTCTGTGCGGGTACTGGACGAGACCCCGGGTA 827
| | | | |
QY 781 GGGGCTGTGTGGGAGCGTGGGATCTTAACCTCTTGGACTCTCTGGGTGAGCTTCCTTA 840
| | | | |
Db 828 GGGGCTGTGTGGGAGCGTGGGATCTTAACCTCTTGGACTCTCTGGGTGAGCTTCCTTA 887
| | | | |
QY 841 AGACGCTGCGATACTGGGTAGCTTCATATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
| | | | |
Db 888 AGACGCTGCGATACTGGGTAGCTTCATATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 947
| | | | |
QY 901 TGTGTATCTGCACTCAGCGGCTCTTAGGCTATATAGGAATAACCTCTTGCAGCAGTTT 960
| | | | |
Db 948 TGTGTATCTGCACTCAGCGGCTCTTAGGCTATATAGGAATAACCTCTTGCAGCAGTTT 1007
| | | | |
QY 961 CTTGAGGCTGTGTACAAACAGAGTTCTCTGGGTACCCACTGTACTTTTGGGATGACCGG 1020
| | | | |
Db 1008 CTTGAGGCTGTGTACAAACAGAGTTCTCTGGGTACCCACTGTACTTTTGGGATGACCGG 1067
| | | | |
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGGTATGTATACCAAGTA 1064
| | | | |
Db 1068 CACCTCACCAACCGCATGCTCAGCATGGGTATGTATACCAAGTA 1111
| | | | |

RESULT 9

US-10-042-523-1

; Sequence 1, Application US/10042523

; GENERAL INFORMATION:

; APPLICANT: Briskin, Michael J.

; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 19-Oct-2001
APPLICATION NUMBER: US/10/042,523
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/635,552A
FILING DATE: 22-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LK995-07
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2116 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 36..1769
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-042-523-1

Query Match 99.9%; Score 1064; DB 46; Length 2116;
Best Local Similarity 100.0%; Pred. No. 2.8e-164;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAGCAGGAGCGCCAGCCCACTCTGACGCGCGCGCTGCTCGGCGTGGCC 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 36 ATGACAGCAGGAGCGCCAGCCCACTCTGACGCGCGCGCTGCTCGGCGTGGCC 95
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 CGGAGGGTGTGACCATCGCTTCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 96 CGGAGGGTGTGACCATCGCTTCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 155
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 GCGCGGGGTGCGGCTCGGCTCCGATCGCTACGGGCTCTGGGCTTACGGG 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 156 GCGCGGGGTGCGGCTCGGCTCCGATCGCTACGGGCTCTGGGCTTACGGG 215
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 GCCTTCTTTCAGCGCACCTGTGCGCAGAGCCTCTTGGGTACTCTGGAGCAGCGCGG 240
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 216 GCCTTCTTTCAGCGCACCTGTGCGCAGAGCCTCTTGGGTACTCTGGAGCAGCGCGG 275
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 GTGGCGGGCGGCGCGGCGCGCTGGATGAGCCACCGCGCGAGTGTGGCGTGAAC 300
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 276 GTGGCGGGCGGCGCGGCGCGCTGGATGAGCCACCGCGCGAGTGTGGCGTGAAC 335
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 ATCTCCGCTTACAGAGAGACCCCGGCTACCTGCGCCAGTGCTGGGCTGCGCGCGCC 360
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 336 ATCTCCGCTTACAGAGAGACCCCGGCTACCTGCGCCAGTGCTGGGCTGCGCGCGCC 395
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 CTGCTGTACCGCGCGCGGCTGCGCTCTCATGTGTGGATGGCAACCGCGCGCAG 420
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 396 CTGCTGTACCGCGCGCGGCTGCGCTCTCATGTGTGGATGGCAACCGCGCGCAG 455
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 GACCTCTACATGGTGCATGTTCCGCGAGGTCTTCGCTGACGAGAACCCCGCCACAGTAC 480
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 456 GACCTCTACATGGTGCATGTTCCGCGAGGTCTTCGCTGACGAGAACCCCGCCACAGTAC 515
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 GTGTGGGACGGCACTACACAGCCCTGGGAAACCCCGCGCGCGCGGCTGGGCGCC 540
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 516 GTGTGGGACGGCACTACACAGCCCTGGGAAACCCCGCGCGCGCGGCTGGGCGCC 575
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 541 GGAGCCTATGGGAGGTGAGCGGAGGATCTCTGGGCGGCTGGCAGTGGAGGCGCTGTG 600

DB 576 GGAGCCTATCGGAGGTGAGCGGAGGATCTTGGCGGCTGGAGTGGAGCGCTGTGTG 635
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 601 AGGACTCCAGGTGCGTGTGCGCGCAGCGCTGGGGCGGCAAGCGGAGGTCAATGTAC 660
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 636 AGGACTCCAGGTGCGTGTGCGCGCAGCGCTGGGGCGGCAAGCGGAGGTCAATGTAC 695
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 661 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGGACTACGTGAGGCTGTGATCGGACACA 720
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 696 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGGACTACGTGAGGCTGTGATCGGACACA 755
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 721 AGGTTGGACCCCATGGCACTCTGCGGGGTACTGGAGAGAGACCCCGGGTA 780
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 756 AGGTTGGACCCCATGGCACTCTGCGGGGTACTGGAGAGAGACCCCGGGTA 815
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 781 GGGGCTGTGTGGGAGCGTCCGATCCTTAACCTCTGGACTCTGCGGTGAGCTTCTTA 840
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 816 GGGGCTGTGTGGGAGCGTCCGATCCTTAACCTCTGGACTCTGCGGTGAGCTTCTTA 875
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 841 AGCAGCCTCGATACTGGGTAGCTTCAATCTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 876 AGCAGCCTCGATACTGGGTAGCTTCAATCTGGAGCGGCTTGTGAGAGCTACTTCCAC 935
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 901 TGTGTATCTGCATCAGCGGTCTCTAGGCTTATAGGAATAACCTCTTGAGCAGGTTT 960
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 936 TGTGTATCTGCATCAGCGGTCTCTAGGCTTATAGGAATAACCTCTTGAGCAGGTTT 995
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 961 CTTGAGGCGTGTACAACAGAGTTCTGCGGTACCCACTACTTTTGGGATGACCGG 1020
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 996 CTTGAGGCGTGTACAACAGAGTTCTGCGGTACCCACTACTTTTGGGATGACCGG 1055
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1021 CACTCACCAACCGCATGCTCAGCATGGGTTATGCTACCAAGTA 1064
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1056 CACTCACCAACCGCATGCTCAGCATGGGTTATGCTACCAAGTA 1099
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 10

PCT-US02-41225A-224
; GENERAL INFORMATION:
; SEQUENCE 224, Application PC/TUS0241225A
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: JONES, Karen Anne
; APPLICANT: SCHAFER, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GR-0022 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/41225A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/342,603
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 224
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: HASI_cds.1
PCT-US02-41225A-224

Query Match 99.8%; Score 1062.4; DB 1; Length 1737;
Best Local Similarity 99.9%; Pred. No. 5.2e-164;
Matches 1063; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACAGCAGGAGCGGCGCCCAAGCCCACTCTGACGCGCGCGCTGCTCGGCGTGGCC 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 CGGAGGGTGTGACCATCGCTTCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 CGGAGGGTGTGACCATCGCTTCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 GCGCGGGGTGCGGCTGCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 180

Db 841 ACAGCGCTGATGAGGATGCTTCAATGTGGAGCGGCTTGTACAGAGCTACTTCCAC 900
QY 901 TGTGTATCTGTGATCAGCGGTCTCTAGGCTTATATAGGAATAAAGCTCTTGTGAGCAGTTT 960
Db 901 TGTGTATCTGTGATCAGCGGTCTCTAGGCTTATATAGGAATAAAGCTCTTGTGAGCAGTTT 960
QY 961 CTTGAGGCTGTGTACAAACAGAGTTTCTTGGGTACCCACTGTACTTTTGGGGATACCGG 1020
Db 961 CTTGAGGCTGTGTACAAACAGAGTTTCTTGGGTACCCACTGTACTTTTGGGGATACCGG 1020
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTATGTACCAAGTA 1064
Db 1021 CACCTCACCAACCGCATGCTCAGCATGGTATGTACCAAGTA 1064

RESULT 12
PCT-US02-41225A-226
; Sequence 226, Application PC/TUS0241225A
; GENERAL INFORMATION:
; APPLICANT: JONES, Karen Anne
; APPLICANT: SCHAFER, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GX-0022 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/41225A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/342,603
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 226
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: HAS1_mrna_build.1
PCT-US02-41225A-226

Query Match 99.8%; Score 1062.4; DB 1; Length 2087;
Best Local Similarity 99.9%; Pred. No. 5.1e-164;
Matches 1063; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGACGACGAGCGCGCCAGCCCACTCTCTGAGCGCGCGTCTCGGCTGCGC 60
Db 36 ATGAGACGACGAGCGCGCCAGCCCACTCTCTGAGCGCGCGTCTCGGCTGCGC 95
QY 61 CGAGGGGTGCTGACCATCGCCTTCCGCTCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db 96 CGAGGGGTGCTGACCATCGCCTTCCGCTCTCATCTGGGCTCATGACCTGGGCTTAC 155
QY 121 GCGCGGGGTGCGCTGCGCTCCGATCGCTACGGCCTCTGGCCTCTGGCCTCTACGG 180
Db 156 GCGCGGGGTGCGCTGCGCTCCGATCGCTACGGCCTCTGGCCTCTACGG 215
QY 181 GCCTTCCTTTGAGCGCATCTGTGTGGCGAGAGCTCTTGGCTACTGTGAGCAGCGCGG 240
Db 216 GCCTTCCTTTGAGCGCATCTGTGTGGCGAGAGCTCTTGGCTACTGTGAGCAGCGCGG 275
QY 241 GTGGCGGGCGCGCGGGGCGCTGTGATGAGCAGCACCGCGCAGTGTGGCTGACC 300
Db 276 GTGGCGGGCGCGCGGGGCGCTGTGATGAGCAGCACCGCGCAGTGTGGCTGACC 335
QY 301 ATCTCCGCTTACAGAGAGACCCCGCTACCTTGCGCCAGTGGCTGGCGTCCGCGCGCC 360
Db 336 ATCTCCGCTTACAGAGAGACCCCGCTACCTTGCGCCAGTGGCTGGCGTCCGCGCGCC 395
QY 361 CTGCTGTACCCCGCGCGGGTGTGGCTCTCATGTGTGGATGGCAACCGCGCCAG 420
Db 396 CTGCTGTACCCCGCGCGGGTGTGGCTCTCATGTGTGGATGGCAACCGCGCCAG 455
QY 421 GACCTCTACATGGTTCGATGTTCCGCGAGGTCTTTCGTGACGAGGACCCCGCCACGTAC 480

Db 456 GACCTCTACATGGTTCGATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCCACGTAC 515
QY 481 GTGTGGAGCGCACTACCAAGCCCTGGGAAACCCGCGGCGCGCGCGCGTGGCGCC 540
Db 516 GTGTGGAGCGCACTACCAAGCCCTGGGAAACCCGCGGCGCGCGCGTGGCGCC 575
QY 541 GGAGCTATCGGAGGTGGAGCGGAGGATCCTTGGGCGGCTGGCAGTGGAGCGCGCTGGTG 600
Db 576 GGAGCTATCGGAGGTGGAGCGGAGGATCCTTGGGCGGCTGGCAGTGGAGCGCGCTGGTG 635
QY 601 AGGACTCGCAGGTGCTGTGGCGCGCAGCGCTGGGCGCGCAAGCGCGAGGTTCATGTAC 660
Db 636 AGGACTCGCAGGTGCTGTGGCGCGCAGCGCTGGGCGCGCAAGCGCGAGGTTCATGTAC 695
QY 661 ACAGCTTTCAAGCGCTCGGAGATTTCGCTGAGCTACGTGACGCTGTGTGACTCGGACACA 720
Db 696 ACAGCTTTCAAGCGCTCGGAGATTTCGCTGAGCTACGTGACGCTGTGTGACTCGGACACA 755
QY 721 AGGTTGGACCCCATGGCACTGCTGAGCTCGTGGGCTTCTGGAGAGAGAGAGAGAGAGAG 780
Db 756 AGGTTGGACCCCATGGCACTGCTGAGCTCGTGGGCTTCTGGAGAGAGAGAGAGAGAGAG 815
QY 781 GGGGCTGTGTGGTGGGAGCGTGGGATCCTTAACCTCTTGGACTCTCTGGGTGAGTTCCTA 840
Db 816 GGGGCTGTGTGGTGGGAGCGTGGGATCCTTAACCTCTTGGACTCTCTGGGTGAGTTCCTA 875
QY 841 AGCAGCTGCGATACCTGCTGAGTTCCTTAAGTGGAGCGGCTTGTGAGAGTACTTCCAC 900
Db 876 AGCAGCTGCGATACCTGCTGAGTTCCTTAAGTGGAGCGGCTTGTGAGAGTACTTCCAC 935
QY 901 TGTGTATCTGATCAGCGGCTCTAGGCTTATAGGAATAAAGCTCTTGGAGTACCTCTTGGAGT 960
Db 936 TGTGTATCTGATCAGCGGCTCTAGGCTTATAGGAATAAAGCTCTTGGAGTACCTCTTGGAGT 995
QY 961 CTTGAGGCTGGTACAAACAGAGTTTCTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
Db 996 CTTGAGGCTGGTACAAACAGAGTTTCTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1055
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTATGTACCAAGTA 1064
Db 1056 CACCTCACCAACCGCATGCTCAGCATGGTATGTACCAAGTA 1099

RESULT 13
US-60-537-226
; Sequence 226, Application US/60257537
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen
; APPLICANT: Schaffer, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GX-0022 P
; CURRENT APPLICATION NUMBER: US/60/257,537
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 226
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: HAS1_mrna_build.1
US-60-537-226

Query Match 99.8%; Score 1062.4; DB 95; Length 2087;
Best Local Similarity 99.9%; Pred. No. 5.1e-164;
Matches 1063; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGACGACGAGCGCGCCAGCCCACTCTCTGAGCGCGCGCTGCTCGGCTGCGC 60
Db 36 ATGAGACGACGAGCGCGCCAGCCCACTCTCTGAGCGCGCGCTGCTCGGCTGCGC 95

QY 841 AGCAGCTGCGATACATGGGTGAGCGGCTTCAATGTGGAGCGGCTTGTGACAGAGCTACTTCCAC 900
Db |||||
QY 876 AGCAGCTGCGATACATGGGTGAGCGGCTTCAATGTGGAGCGGCTTGTGACAGAGCTACTTCCAC 935
Db |||||
QY 901 TGTGTATCTGTCATCAGCGGCTCTTAGCCCTATATAGGAATAAATCTTGTGACAGAGTTT 960
Db |||||
QY 936 TGTGTATCTGTCATCAGCGGCTCTTAGCCCTATATAGGAATAAATCTTGTGACAGAGTTT 995
Db |||||
QY 961 CTTGAGGCTGTGTACAAACAGAGTTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
Db |||||
QY 996 CTTGAGGCTGTGTACAAACAGAGTTCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1055
Db |||||
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1064
Db |||||
QY 1056 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1099
Db |||||
RESULT 15
US-10-940-774-5659
; Sequence 5659, Application US/10940774
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/10/940,774
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Human
US-10-940-774-5659
Query Match 99.6%; Score 1060.8; DB 66; Length 2087;
Best Local Similarity 99.8%; Pred. No. 9.3e-164;
Matches 1062; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCCTGACAGCCCGCGCTGTCTCGGCGCTGACC 60
Db |||||
QY 61 CGGAGGTGTGACCATCGCTTCGCCCTGCTCATCTCGGCGCTCATGACCTGGGCGCTTAC 120
Db |||||
QY 96 CGGAGGTGTGACCATCGCTTCGCCCTGCTCATCTCGGCGCTCATGACCTGGGCGCTTAC 155
Db |||||
QY 121 CGCGCGGGGTGCGCGCTCGGCTCGCATCGCTACGGCCCTCTCGGCGCTTACCGG 180
Db |||||
QY 156 CGCGCGGGGTGCGCGCTCGGCTCGCATCGCTACGGCCCTCTCGGCGCTTACCGG 215
Db |||||
QY 181 GCCTTCCTTTTACGCGCACCTGGTGGCGCAGAGCTTTTCGGCTACTTGAGACACCGGCGG 240
Db |||||
QY 216 GCCTTCCTTTTACGCGCACCTGGTGGCGCAGAGCTTTTCGGCTACTTGAGACACCGGCGG 275
Db |||||
QY 241 GTGGGCGGGCGGCGGCGGCGGCGCTGATGATGACGACCGCGCGCAGTGTGGCGCTGACC 300
Db |||||
QY 276 GTGGGCGGGCGGCGGCGGCGGCGCTGATGATGACGACCGCGCGCAGTGTGGCGCTGACC 335
Db |||||
QY 301 ATCTCCGCTTACGAGGAGACCGCGGTACTCTGGCGCAGTGCCTGGCGCTCGCGCGCGCC 360
Db |||||
QY 336 ATCTCCGCTTACGAGGAGACCGCGGTACTCTGGCGCAGTGCCTGGCGCTCGCGCGCGCC 395
Db |||||
QY 361 CTGCTGTATACCGCGCGCGCGCTGCGGCTCTCATGTGTGGATGGAACCGCGCGGAG 420
Db |||||
QY 396 CTGCTGTATACCGCGCGCGCGCTGCGGCTCTCATGTGTGGATGGAACCGCGCGGAG 455
Db |||||

Search completed: March 13, 2005, 16:17:29
Job time : 5178.05 secs

QY 421 GACCTCTACATGTGTCGACATGTTCCGCGAGGTCTTCCGTGACGAGGACCCCGCCACGTCAC 480
Db |||||
QY 456 GACCTCTACATGTGTCGACATGTTCCGCGAGGTCTTCCGTGACGAGGACCCCGCCACGTCAC 515
Db |||||
QY 481 GTGTGGAGCGGCAACTACACAGCCCTTGGGAACCCCGCGCGGCGCGGTGTGGCGGCC 540
Db |||||
QY 516 GTGTGGAGCGGCAACTACACAGCCCTTGGGAACCCCGCGCGGCGCGGTGTGGCGGCC 575
Db |||||
QY 541 GGAGCCTATCGGAGGTGAGGCGGAGAGATCTTGGGCGGCTGGCAGTGTGGAGGCGCTGGTG 600
Db |||||
QY 576 GGAGCCTATCGGAGGTGAGGCGGAGAGATCTTGGGCGGCTGGCAGTGTGGAGGCGCTGGTG 635
Db |||||
QY 601 AGGACTCGCAGGTGCGTGTGCGCGCAGCGCTGCGGCGGCAAGCGCGAGGTCTATGTAC 660
Db |||||
QY 636 AGGACTCGCAGGTGCGTGTGCGCGCAGCGCTGCGGCGGCAAGCGCGAGGTCTATGTAC 695
Db |||||
QY 661 ACAGCCTTCAAGCGCTCGGAGATTTCGGTGGACTACGTGACGGTCTGTGACTCGGACACA 720
Db |||||
QY 696 ACAGCCTTCAAGCGCTCGGAGATTTCGGTGGACTACGTGACGGTCTGTGACTCGGACACA 755
Db |||||
QY 721 AGGTGGACCCCATGGCACCTGCTGGAGCTGCTGGGCTACTGGACGAGGACCCCGGGTA 780
Db |||||
QY 756 AGGTGGACCCCATGGCACCTGCTGGAGCTGCTGGGCTACTGGACGAGGACCCCGGGTA 815
Db |||||
QY 781 GGGGCTGTGTGGGGGACGTGCGGATCCTTAACCTCTGGACTCCTGGGTGAGCTTCTTA 840
Db |||||
QY 816 GGGGCTGTGTGGGGGATGTGGGATCCTTAACCTCTGGACTCCTGGGTGAGCTTCTTA 875
Db |||||
QY 841 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGACAGCTACTTCCAC 900
Db |||||
QY 876 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGACAGCTACTTCCAC 935
Db |||||
QY 901 TGTGTATCTGTCATCAGCGGCTCTTAGGCTATATAGGCAATAACCTCTTGCAGAGTTT 960
Db |||||
QY 936 TGTGTATCTGTCATCAGCGGCTCTTAGGCTATATAGGCAATAACCTCTTGCAGAGTTT 995
Db |||||
QY 961 CTTGAGGCGCTGTGTACCAACGAGGTTCTCGGTTACCCACTGTACTTTTGGGGATGACCGG 1020
Db |||||
QY 996 CTTGAGGCGCTGTGTACCAACGAGGTTCTCGGTTACCCACTGTACTTTTGGGGATGACCGG 1055
Db |||||
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1064
Db |||||
QY 1056 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1099
Db |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:46:32 ; Search time 90 Seconds
(without alignments)
1375.148 Million cell updates/sec

Title: US-10-672-399-8

Perfect score: 1697

Sequence: 1 AFLSAHLVAQSLFAYLEHRR.....SGTPGPAATQRRPRPSCGG 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980a:*
- 2: geneseqp1990a:*
- 3: geneseqp2000a:*
- 4: geneseqp2001a:*
- 5: geneseqp2002a:*
- 6: geneseqp2003a:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1553.5	91.5	578	2	Aaw26765 Human hya
2	1546.5	91.1	543	2	Aaw36503 Human hya
3	1508.5	88.9	582	3	Aay78135 Recombina
4	1508.5	88.9	583	2	Aaw30704 Mouse hya
5	1508.5	88.9	583	2	Aay32503 Hyaluroa
6	1508.5	88.9	583	3	Aay68491 Mouse hya
7	1508.5	88.9	583	3	Aab09948 Murine HA
8	1508.5	88.9	583	6	ABP96028 Mouse hya
9	1508.5	88.9	583	7	Aae39152 Mouse hya
10	1458.5	85.9	563	3	Aay78129 Recombina
11	1362	80.3	584	3	Aay78132 Recombina
12	1358.5	80.1	582	3	Aay78138 Recombina
13	1358.5	80.1	583	3	Aay78131 Recombina
14	1085	63.9	552	3	Aay78133 Recombina
15	1085	63.9	553	3	Aay78136 Recombina
16	965.5	56.9	588	7	Add93328 Xenopus 1
17	948.5	55.9	552	7	ADC49213 Rabbit hy
18	947	55.8	552	3	Aay78142 Recombina
19	943	55.6	577	3	Aay78128 Recombina
20	942	55.5	552	3	Aay78140 Recombina
21	942	55.5	552	3	Aay78139 Recombina
22	938.5	55.3	553	3	Aay78134 Recombina
23	935.5	55.1	554	2	Aaw50010 Murine hy
24	935.5	55.1	554	3	Aay68493 Mouse hya
25	935.5	55.1	554	6	ABP96030 Mouse hya

26	935.5	55.1	554	7	AAE39154	Mouse hya
27	935.5	55.1	554	8	ADJ76214	Marker ge
28	935.5	55.1	557	3	AAY78130	Recombina
29	935	55.1	552	2	AAW50009	Murine hy
30	935	55.1	552	3	AAY68492	Mouse hya
31	935	55.1	552	5	ABB57265	Mouse 1sc
32	935	55.1	552	6	ABP96029	Mouse hya
33	935	55.1	552	7	ADA49686	Mouse hya
34	935	55.1	552	7	AAE39153	Mouse hya
35	935	55.1	552	7	ADC59299	Mouse HAS
36	935	55.1	553	3	AAY78137	Recombina
37	934	55.0	552	7	ADA49684	Human hya
38	934	55.0	552	7	ADC59297	Human HAS
39	934	55.0	552	7	ADD48818	Human Pro
40	931.5	54.9	573	3	AAY78127	Recombina
41	930	54.8	552	3	AAW78141	Recombina
42	930	54.8	552	7	ADD48816	Rat Prote
43	928	54.7	553	6	ABR48209	Human bla
44	928	54.7	553	6	ABU56490	Lung canc
45	928	54.7	553	6	ABU56686	Lung canc

ALIGNMENTS

RESULT 1

AAW26765

ID AAW26765 standard; protein; 578 AA.

AC AAW26765;

DT 21-MAY-1998 (first entry)

DE Human hyaluronan synthase.

XX Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
KW wound healing; vulnery; tissue repair; scar; keloid; therapy.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Binding-site 119..127

FT Modified-site /note= "putative hyaluronan binding site"

FT Modified-site 340..343

FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"

FT Modified-site 354..356

FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"

FT Modified-site 371..374

FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"

FT Modified-site 378..380

FT Modified-site /note= "consensus phosphorylation sequence for cAMP-dependent protein kinase"

XX WO9740174-A1.

PN 30-OCT-1997.

XX 17-APR-1997; 97WO-US006350.

XX 22-APR-1996; 96US-00635552.

XX (LEUK-) LEUKOSITE INC.

XX Briekin MJ;

XX WPI; 1997-549359/50.

XX N-PSDB; AAT99541.

XX Human hyaluronan synthase - useful for recombinant production of

PT hyaluronan acid for wound healing, tissue repair and reducing

PT hypertrophic scar and keloid formation.

PS Claim 5; Page 38-40; 58pp; English.

XX This protein comprises human hyaluronan synthase (HAS), an enzyme involved in the synthesis of hyaluronan (hyaluronic acid) and which has the ability to confer cell adhesion by the lymphocyte receptor CD44. The amino acid sequence was deduced from clone 30C (see AAT99541), which was obtained from a human mesenteric lymph node expression library using an expression cloning system developed to isolate cDNA clones that encode proteins that confer adhesion of the murine T cell lymphoma TK1. The isolated clone can be utilised in a claimed method for producing HAS in host cells. Such host cells are used in a claimed method for the production of hyaluronan. Hyaluronan is useful for wound healing and tissue repair, and can reduce or prevent hypertrophic scars and keloid formation. It is also used in eye surgery as a replacement for vitreous fluid

XX Sequence 578 AA;

Query Match 91.5%; Score 1553.5; DB 2; Length 578;
Best Local Similarity 93.1%; Pred. No. 1.8e-172;
Matches 298; Conservative 2; Mismatches 1; Indels 19; Gaps 1;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
DB |||||
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB |||||
QY 61 LLYPARLRVLMVVDGNRAEDLYMDFREVPFADDPATYWDGNYHQPWPAAAGAVGA 120
DB |||||
QY 121 LLYPARLRVLMVVDGNRAEDLYMDFREVPFADDPATYWDGNYHQPWPAAAGAVGA 180
DB |||||
QY 121 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTFKALGDSVDYVQVCDSDT 180
DB |||||
QY 181 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTFKALGDSVDYVQVCDSDT 240
DB |||||
QY 181 RLDPMALLELRVLDDEPRVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 240
DB |||||
QY 241 RLDPMALLELRVLDDEPRVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
DB |||||
QY 241 CVSCISGPLGLYRNLLQOFLQFLYAWNQKFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
DB 301 CVSCISGPLGLYRNLLQOFLQFLYAWNQKFLGTHCTFGDDRHLTNRMLSMGYATKTSRSR 360
QY 295 -----AEGTRWS 301
DB 361 CYSETPSSFLRWLSQQTRWS 380

RESULT 2
AAW36503
ID AAW36503 standard; protein; 543 AA.

XX AAW36503;

XX 22-APR-1998 (first entry)

XX Human hyaluronate synthetase.

XX Hyaluronate synthetase; human; hyaluronic acid; drug preparation;
KW cosmetic preparation; gene therapy; carcinogenesis.

XX Homo sapiens.

XX WO9738113-A1.

XX 16-OCT-1997.

XX 31-MAR-1997; 97WO-JP001111.

XX 05-APR-1996; 96JP-00084326.

XX 30-APR-1996; 96JP-00109663.

XX

PA (SEKG) SEIKAGAKU CORP.

PI Itano N, Kimata K;

XX WPI; 1997-512726/47.

DR N-PSDB; AAT96713.

XX DNA encoding human hyaluronate synthetase - for industrial scale
PT production of hyaluronic acid used in generating anti-carcinogenic drugs
or for cosmetics.

XX Claim 2; Page 23-27; 35pp; Japanese.

XX This sequence represents a human hyaluronate synthetase, and is encoded
CC by the coding sequence of the invention. This enzyme is useful for
CC industrial scale production of hyaluronic acid for use in the preparation
CC of drugs and cosmetics. The drugs can also be used in compositions for
CC the treatment of disorders involving the lowering of hyaluronic acid
CC production. The peptides may be used for the preparation of antibodies
CC recognising hyaluronate synthetase, e.g. for diagnostic purposes.
CC Antisense DNA or RNA corresponding to the DNA can be used in gene therapy
CC treatment of carcinogenesis

XX Sequence 543 AA;

Query Match 91.1%; Score 1546.5; DB 2; Length 543;
Best Local Similarity 92.5%; Pred. No. 1.1e-171;
Matches 296; Conservative 4; Mismatches 1; Indels 19; Gaps 1;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
DB |||||
QY 26 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 85
DB |||||
QY 61 LLYPARLRVLMVVDGNRAEDLYMDFREVPFADDPATYWDGNYHQPWPAAAGAVGA 120
DB |||||
QY 86 LLYPARLRVLMVVDGNRAEDLYMDFREVPFADDPATYWDGNYHQPWPAAAGAVGA 145
DB |||||
QY 121 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTFKALGDSVDYVQVCDSDT 180
DB |||||
QY 146 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTFKALGDSVDYVQVCDSDT 205
DB |||||
QY 181 RLDPMALLELRVLDDEPRVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 240
DB |||||
QY 206 RLDPMALLELRVLDDEPRVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 265
DB |||||
QY 241 CVSCISGPLGLYRNLLQOFLQFLYAWNQKFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
DB 266 CVSCISGPLGLYRNLLQOFLQFLYAWNQKFLGTHCTFGDDRHLTNRMLSMGYATKTSRSR 325
QY 295 -----AEGTRWS 301
DB 326 CYSETPSSFLRWLSQQTRWS 345

RESULT 3

AAW78135
ID AAW78135 standard; protein; 582 AA.

XX AAW78135;

XX 27-APR-2000 (first entry)

XX Recombinant chimeric hyaluronate synthase modified protein #9.

XX Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.

XX Mus sp.

XX Synthetic.

XX Chimeric.

XX JP2000004886-A.

XX

DT 19-OCT-1999 (first entry)
 XX Hyaluronate synthase protein sequence.
 XX Hyaluronate synthase; promoter; mouse; cell growth inhibitor.

OS Mus sp.
 PN JP11196875-A.
 XX 27-JUL-1999.
 XX 14-JAN-1998; 98JP-00006191.
 XX 14-JAN-1998; 98JP-00006191.
 XX (SEKG) SEIKAGAKU KOGYO CO LTD.

XX WPI; 1999-496653/42.
 DR N-PSDB; AAZ10862.

XX New promoter DNA of hyaluronate synthase - used to specifically express
 PT gene with cell growth inhibiting activity.

XX Example 1; Page 8-10; 13pp; Japanese.

XX This sequence is encoded by the mouse hyaluronate synthase gene, which
 CC was used to isolate the hyaluronate synthase promoter of the invention.
 CC The promoter can be used for specifically expressing a gene having cell
 CC growth inhibiting activity

XX Sequence 583 AA;

Query Match 88.9%; Score 1508.5; DB 2; Length 583;
 Best Local Similarity 89.9%; Pred. No. 3.6e-167;
 Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
 DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119

QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPA-A 114
 DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPAE 179

QY 115 AGAVGAGAYREVEAEDPGRLAVALVTRRCVCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 174
 DB 180 TGAVGEGAYREVEAEDPGRLAVALVTRRCVCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 239

QY 175 VCDSDTRLDPMALLELVRLVLDDEPRVGAAGGVDRILNPLDSWVSLSLRYWVAFNVERA 234
 DB 240 VCDSDTRLDPMALLELVRLVLDDEPRVGAAGGVDRILNPLDSWVSLSLRYWVAFNVERA 299

QY 235 CQSYFHCVCISGPGLYRNLLQOFLQFLEAWYNQKFLGTHCTFGDDRHLTNRLMSMGYATK 294
 DB 300 CQSYFHCVCISGPGLYRNLLQOFLQFLEAWYNQKFLGTHCTFGDDRHLTNRLMSMGYATK 359

QY 295 -----AEGTRWS 301
 DB 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 6
 AAY68491
 ID AAY68491 standard; protein; 583 AA.
 XX AC AAY68491;
 XX 27-APR-2000 (first entry)
 DT Mouse hyaluronate synthase modified protein HAS1.
 DE Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
 KW

research reagent; biochemical research; medical development.

Mus sp.

JP2000004886-A.

11-JAN-2000.

24-JUN-1998; 98JP-00193788.

24-JUN-1998; 98JP-00193788.

(SEKG) SEIKAGAKU KOGYO CO LTD.

WPI; 2000-140125/13.

N-PSDB; AAZ88199.

A hyaluronate synthase modified protein - useful as a research reagent
 for biochemical research and medical development.

Claim 1; Page 14-16; 30pp; Japanese.

The present invention describes a recombinant protein which consists of
 three continuous regions (N-terminal region, internal region and C-
 terminal region) where one or two regions among the above three regions
 is selected from the three hyaluronate synthase (HAS) modified proteins
 HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
 HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
 is useful as a research reagent for biochemical research and medical
 development. The invention provides a HAS modified protein of increased
 or lowered activity. The present sequence represents mouse HAS1

Sequence 583 AA;

Query Match 88.9%; Score 1508.5; DB 3; Length 583;
 Best Local Similarity 89.9%; Pred. No. 3.6e-167;
 Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
 DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119

QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPA-A 114
 DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPAE 179

QY 115 AGAVGAGAYREVEAEDPGRLAVALVTRRCVCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 174
 DB 180 TGAVGEGAYREVEAEDPGRLAVALVTRRCVCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 239

QY 175 VCDSDTRLDPMALLELVRLVLDDEPRVGAAGGVDRILNPLDSWVSLSLRYWVAFNVERA 234
 DB 240 VCDSDTRLDPMALLELVRLVLDDEPRVGAAGGVDRILNPLDSWVSLSLRYWVAFNVERA 299

QY 235 CQSYFHCVCISGPGLYRNLLQOFLQFLEAWYNQKFLGTHCTFGDDRHLTNRLMSMGYATK 294
 DB 300 CQSYFHCVCISGPGLYRNLLQOFLQFLEAWYNQKFLGTHCTFGDDRHLTNRLMSMGYATK 359

QY 295 -----AEGTRWS 301
 DB 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 7
 AAB09948
 ID AAB09948 standard; protein; 583 AA.
 XX AC AAB09948;
 XX 19-OCT-2000 (first entry)
 DT Murine HAS1 protein.
 DE
 XX


```

|||||
Db 240 VCDSDTRLDPALLELVRLVLEDDPRVGAAGVDVRLINPLDSWVSFLSSRLRWAFNVERA 299
QY 235 CQSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLTNRMLSGYATK 294
Db 300 CQSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLTNRMLSGYATK 359
QY 295 -----AEGTRWS 301
Db 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 9
AAE39152
ID AAE39152 standard; protein; 583 AA.
XX
AC AAE39152;
XX
DT 18-DEC-2003 (first entry)
XX
DE Mouse hyaluronan synthase (HAS) 1.
XX
KW Mouse; hyaluronic acid; HA; hyaluronan synthase; HAS; dry eye syndrome;
KW eye cell; osteoarthritis; gene therapy; enzyme.
XX
OS Mus sp.
XX
PN US2003087850-A1.
XX
PD 08-MAY-2003.
XX
PF 10-JUL-2001; 2001US-00902939.
XX
PR 10-JUL-2001; 2001US-00902939.
XX
PA (DEHA/) DEHAZYA P.
XX (CHEN/) CHEN W.
XX
PI Dehazy P, Chen W;
DR N-PSDB; AAD59442.
XX
PT Dihyrazide derivatized hyaluronic acid/nucleic acid bioconjugate
PT comprising derivatized hyaluronic acid cross linked to nucleic acid
PT encoding hyaluronan synthase useful for treating dry eye syndrome.
XX
PS Claim 18; Page 15-16; 31pp; English.
XX
CC The invention relates to dihydrazide derivatised hyaluronic acid (HA)/
CC nucleic acid bioconjugate comprising derivatised HA cross linked to
CC nucleic acid encoding hyaluronan synthase (HAS). The invention is useful
CC for transfecting an eye cell of an individual. It is useful for treating
CC dry eye syndrome and osteoarthritis of the particular joints. The
CC invention is also useful in gene therapy. The present sequence is mouse
CC HAS1 enzyme
XX
SQ Sequence 583 AA;
Query Match 88.9%; Score 1508.5; DB 7; Length 583;
Best Local Similarity 89.9%; Pred. No. 3.6e-167; Indels 25; Gaps 3;
Matches 293; Conservative 2; Mismatches 6;
QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
Db 60 AFLSAHLVAQSLFAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRLRLVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNVHQWPEA-A 114
Db 120 TSARALLYPHTRLRLVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNVHQWPEA 179
QY 115 AGAVGAGAYREVEADPGRLAVEALVRTRRCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 174
|||||

```

```

Db 180 TGAVGEGAYREVEADPGRLAVEALVRTRRCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 239
QY 175 VCDSDTRLDPALLELVRLVLEDDPRVGAAGVDVRLINPLDSWVSFLSSRLRWAFNVERA 234
Db 240 VCDSDTRLDPALLELVRLVLEDDPRVGAAGVDVRLINPLDSWVSFLSSRLRWAFNVERA 299
QY 235 CQSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLTNRMLSGYATK 294
Db 300 CQSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLTNRMLSGYATK 359
QY 295 -----AEGTRWS 301
Db 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 10
AAAY78129
ID AAAY78129 standard; protein; 563 AA.
XX
AC AAAY78129;
XX
DT 27-APR-2000 (first entry)
XX
DE Recombinant chimeric hyaluronate synthase modified protein #3.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX
PN JP2000004886-A.
XX
PD 11-JAN-2000.
XX
PE 24-JUN-1998; 98JP-00193788.
XX
PR 24-JUN-1998; 98JP-00193788.
XX
PA (SEKG ) SEIKAGAKU KOGYO CO LTD.
XX
DR WPI; 2000-140125/13.
XX
PT A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
PS Claim 4; Page; 30pp; Japanese.
XX
CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAAY78127 to AAAY78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim
XX
SQ Sequence 563 AA;
Query Match 85.9%; Score 1458.5; DB 3; Length 563;
Best Local Similarity 89.2%; Pred. No. 2.5e-161;
Matches 282; Conservative 3; Mismatches 6; Indels 25; Gaps 3;
QY 11 SLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db 51 ALFAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCLTSARALLYPH 110
QY 66 ARLRLVLMVDGNRAEDLYWDMFREVFADEDPATYVWDGNVHQWPEA-AAGAVGAGAYR 124
|||||

```

Db 111 TELRLVLMVVDGNAEDLYVMDNFREVFADEDPATYVMDGNYPWEPAEATGAVGEGAYR 170
QY 125 EVEAEDPGRLAVALVTRTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDTRLDLP 184
Db 171 EVEAEDPGRLAVALVTRTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDTRLDLP 230
QY 185 MALLELRVLDDEPRVGAAGVGGDVRILNPLDSWVSFLSSLRVYVAFNVERACOSYFHCYSC 244
Db 231 MALLELRVLDDEPRVGAAGVGGDVRILNPLDSWVSFLSSLRVYVAFNVERACOSYFHCYSC 290
QY 245 TSGPLGLYRNLLQOFLFLEAWYVQKFLGTHCTGDDRHLLTNRLMSMGYATK----- 294
Db 291 TSGPLGLYRNLLQOFLFLEAWYVQKFLGTHCTGDDRHLLTNRLMSMGYATKYSRRCYSE 350
QY 295 -----AEGTRWS 301
Db 351 TPSSFRLWLSQOTRWS 366

RESULT 11
AAV78132
ID AAV78132 standard; protein; 584 AA.
XX
AC AAV78132;
XX
DT 27-APR-2000 (first entry)
XX
DE Recombinant chimeric hyaluronate synthase modified protein #6.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX JP20000004886-A.
XX
XX 11-JAN-2000.
XX
XX 24-JUN-1998; 98JP-00193788.
XX
XX 24-JUN-1998; 98JP-00193788.
XX (SEKG) SEIKAGAKU KOGYO CO LTD.
XX
XX WPI; 2000-140125/13.
XX
XX A hyaluronate synthase modified protein - useful as a research reagent
XX for biochemical research and medical development.
XX
XX Claim 7; Page; 30pp; Japanese.

XX The present invention describes a recombinant protein which consists of
XX three continuous regions (N-terminal region, internal region and C-
XX terminal region) where one or two regions among the above three regions
XX is selected from the three hyaluronate synthase (HAS) modified proteins
XX HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
XX HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
XX is useful as a research reagent for biochemical research and medical
XX development. The invention provides a HAS modified protein of increased
XX or lowered activity. AAV78127 to AAV78142 represent specifically claimed
XX examples of recombinant proteins from the present invention. N.B. The
XX present sequence is not given in the present specification, but is
XX derived from sequences given as specified in the claim

XX Sequence 584 AA;

Query Match 80.3%; Score 1362; DB 3; Length 584;
Best Local Similarity 80.4%; Pred. No. 5.6e-150;
Matches 263; Conservative 19; Mismatches 19; Indels 26; Gaps 4;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
Db 60 AFLSAHLVAQSLFAYLEHRRVAAAARRSLAKGFLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRLRLRVLMVVDGNAEDLYVMDNFREVFADEDPATYVMDGNYPWEPAEPA-A 114
Db 120 TSARALLYPHTRLRLRVLMVVDGNAEDLYVMDNFREVFADEDPATYVMDGNYPWEPABEA 179
QY 115 AGAVGAGAYREVEAEDPGRLAVALVTRTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYV 173
Db 180 TGAVGEGAYREVEAEDPGRLAVALVTRTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYV 239
QY 174 QVCDSDTRLDPMALLELRVLDDEPRVGAAGVGGDVRILNPLDSWVSFLSSLRVYVAFNVER 233
Db 240 QVCDSDTVLDPACTIEMLRVLEEDPQVGVGGDVQILNKYDSWISFLSSVRVYMAFNVER 299
QY 234 ACQSYFHCYSCISGPLGLYRNLLQOFLFLEAWYVQKFLGTHCTGDDRHLLTNRLMSMGYAT 293
Db 300 ACQSYFHCYSCISGPLGLYRNLLQOFLFLEAWYVQKFLGTHCTGDDRHLLTNRLMSMGYAT 359
QY 294 K-----AEGTRWS 301
Db 360 KYTARSKCLTETPTTRYLRWLNQOTRWS 386

RESULT 12
AAV78138
ID AAV78138 standard; protein; 582 AA.
XX
AC AAV78138;
XX
DT 27-APR-2000 (first entry)
XX
DE Recombinant chimeric hyaluronate synthase modified protein #12.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX JP20000004886-A.
XX
XX 11-JAN-2000.
XX
XX 24-JUN-1998; 98JP-00193788.
XX
XX 24-JUN-1998; 98JP-00193788.
XX (SEKG) SEIKAGAKU KOGYO CO LTD.
XX
XX WPI; 2000-140125/13.
XX
XX A hyaluronate synthase modified protein - useful as a research reagent
XX for biochemical research and medical development.
XX
XX Claim 13; Page; 30pp; Japanese.

XX The present invention describes a recombinant protein which consists of
XX three continuous regions (N-terminal region, internal region and C-
XX terminal region) where one or two regions among the above three regions
XX is selected from the three hyaluronate synthase (HAS) modified proteins
XX HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
XX HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
XX is useful as a research reagent for biochemical research and medical
XX development. The invention provides a HAS modified protein of increased
XX or lowered activity. AAV78127 to AAV78142 represent specifically claimed
XX examples of recombinant proteins from the present invention. N.B. The
XX present sequence is not given in the present specification, but is
XX derived from sequences given as specified in the claim

XX Sequence 582 AA;

Query Match 80.1%; Score 1358.5; DB 3; Length 582;
Best Local Similarity 79.4%; Pred. No. 1.4e-149;
Matches 259; Conservative 21; Mismatches 21; Indels 25; Gaps 3;
SQ Sequence 583 AA;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPA-A 114
DB 120 TSARALLYPHRLRLVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPA 179
QY 115 AGAVGAGAYREVEAEDPGLRAVEALVRTRRCVCAQWRGKREVMYTAFAKALGDSVDVYQ 174
DB 180 TGAVGEGAYREVEAEDPGLRAVEALVRTRRCVCAQWRGKREVMYTAFAKALGDSVDVYQ 239
QY 175 VCDSDTRLDPMALLRLVRVLDDEPRVGAAGVDVRLNPLDSVSVFLSSRLRYWVAFNVERA 234
DB 240 VCDSDTMDLPASSVEMVKVLEEDPMVGGVGDVQILNKYDSWISFLSSRVYMAFNIERA 299
QY 235 COSYFHCVCISGPIGLYRNLLQOFLBAYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 294
DB 300 COSYFGVCQCISGPIGLMYRNSLLHEFVEDWYQEFMGNCQSGFGDDRHLTNRVLSLGYATK 359
QY 295 -----AEGTRWS 301
DB 360 YTARSKCLTETPIEYLRWLNQOTRWS 385

RESULT 13
AA78131
ID AA78131 standard; protein; 583 AA.
XX AA78131;
XX
XX 27-APR-2000 (first entry)
XX
XX Recombinant chimeric hyaluronate synthase modified protein #5.
XX
XX Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
XX research reagent; biochemical research; medical development; chimeric.
XX
XX Mus sp.
XX Synthetic.
XX Chimeric.
XX JP2000004886-A.
XX
XX 11-JAN-2000.
XX
XX 24-JUN-1998; 98JP-00193788.
XX
XX 24-JUN-1998; 98JP-00193788.
XX
XX (SEK) SEIKAGAKU KOGYO CO LTD.
XX
XX WPI; 2000-140125/13.
XX
XX A hyaluronate synthase modified protein - useful as a research reagent
XX for biochemical research and medical development.
XX
XX Claim 6; Page; 30pp; Japanese.

The present invention describes a recombinant protein which consists of three continuous regions (N-terminal region, internal region and C-terminal region) where one or two regions among the above three regions is selected from the three hyaluronate synthase (HAS) modified proteins HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71, HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein is useful as a research reagent for biochemical research and medical development. The invention provides a HAS modified protein of increased or lowered activity. AA78127 to AA78142 represent specifically claimed

CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim
XX
SQ Sequence 583 AA;

Query Match 80.1%; Score 1358.5; DB 3; Length 583;
Best Local Similarity 79.4%; Pred. No. 1.4e-149;
Matches 259; Conservative 21; Mismatches 21; Indels 25; Gaps 3;
SQ Sequence 583 AA;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPA-A 114
DB 120 TSARALLYPHRLRLVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPA 179
QY 115 AGAVGAGAYREVEAEDPGLRAVEALVRTRRCVCAQWRGKREVMYTAFAKALGDSVDVYQ 174
DB 180 TGAVGEGAYREVEAEDPGLRAVEALVRTRRCVCAQWRGKREVMYTAFAKALGDSVDVYQ 239
QY 175 VCDSDTRLDPMALLRLVRVLDDEPRVGAAGVDVRLNPLDSVSVFLSSRLRYWVAFNVERA 234
DB 240 VCDSDTMDLPASSVEMVKVLEEDPMVGGVGDVQILNKYDSWISFLSSRVYMAFNIERA 299
QY 235 COSYFHCVCISGPIGLYRNLLQOFLBAYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 294
DB 300 COSYFGVCQCISGPIGLMYRNSLLHEFVEDWYQEFMGNCQSGFGDDRHLTNRVLSLGYATK 359
QY 295 -----AEGTRWS 301
DB 360 YTARSKCLTETPIEYLRWLNQOTRWS 385

RESULT 14
AA78133
ID AA78133 standard; protein; 552 AA.
XX AA78133;
XX
XX 27-APR-2000 (first entry)
XX
XX Recombinant chimeric hyaluronate synthase modified protein #7.
XX
XX Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
XX research reagent; biochemical research; medical development; chimeric.
XX
XX Mus sp.
XX Synthetic.
XX Chimeric.
XX JP2000004886-A.
XX
XX 11-JAN-2000.
XX
XX 24-JUN-1998; 98JP-00193788.
XX
XX 24-JUN-1998; 98JP-00193788.
XX
XX (SEK) SEIKAGAKU KOGYO CO LTD.
XX
XX WPI; 2000-140125/13.

A hyaluronate synthase modified protein - useful as a research reagent for biochemical research and medical development.
Claim 8; Page; 30pp; Japanese.
The present invention describes a recombinant protein which consists of three continuous regions (N-terminal region, internal region and C-terminal region) where one or two regions among the above three regions is selected from the three hyaluronate synthase (HAS) modified proteins

PS Claim 11; Page: 30pp; Japanese.

Search completed: March 11, 2005, 14:17:16
Job time : 91 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:06:48 ; Search time 18.7879 Seconds
(without alignments)
1638.789 Million cell updates/sec

Title: US-10-672-399-8
Perfect score: 1697
Sequence: 1 AFLSAHLVAQSLFAYLEHRR.....SGTPGPAATORRRPSCGG 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1550.5	91.4	543	2 JC4812	hyaluronan synthas
2	965.5	56.9	588	2 A43740	DG42 protein - Afr
3	409	24.1	568	2 T17588	hyaluronoglucosami
4	292	17.2	424	1 ZZ2RCL	nodulation protein
5	279	16.4	426	2 A95321	NodC N-ACETYLGLUCO
6	276	16.3	424	1 ZZ2RC4	nodulation protein
7	261	15.4	424	1 S12793	nodulation protein
8	257.5	15.2	413	1 S34305	nodulation protein
9	226.5	13.3	395	1 JQ3396	nodulation protein
10	214	12.6	366	2 E59102	hypothetical prote
11	213.5	12.6	395	2 A48755	hyaluronan synthas
12	213.5	12.6	419	2 A53100	hyaluronate syntha
13	185	10.9	1086	2 JC6079	chitin synthase (E
14	177	10.4	1195	2 S61886	chitin synthase (E
15	174	10.3	1165	1 S45879	chitin synthase (E
16	172	10.1	1112	2 T30202	probable chitin sy
17	168	9.9	1041	2 T31097	chitin synthase (E
18	167.5	9.9	1239	2 T42020	class IV chitin sy
19	160	9.4	417	2 JC4547	polysaccharide syn
20	158	9.3	428	1 E38180	nodulation protein
21	157	9.3	414	2 S18962	FBF15 protein - St
22	156.5	9.2	1175	2 S39951	chitin synthase (E
23	148.5	8.8	1103	2 T42022	probable chitin sy
24	143	8.4	1869	2 A59290	class V chitin syn
25	136	8.0	1498	2 S78102	chitin synthase (E
26	133	7.8	1852	2 JC5546	chitin synthase (E
27	132	7.8	447	2 A97211	glycosyltransferas
28	127	7.5	743	2 T34632	probable bi-functi
29	124	7.3	1380	2 T25284	hypothetical prote

30	119	7.0	482	2 G83928	hypothetical prote
31	117	6.9	420	2 D69769	cellulose synthase
32	113	6.7	1009	2 S20538	chitin synthase (E
33	112	6.6	420	2 AG2473	hypothetical prote
34	111.5	6.6	412	2 B90075	intercellular adhe
35	106	6.2	352	2 F75099	rhannosyl transfer
36	104.5	6.2	412	2 S77608	probable intercell
37	100.5	5.9	425	2 F97108	probable glycosylt
38	100	5.9	1013	2 JC2314	chitin synthase (E
39	99.5	5.9	1155	2 AC2675	chromosome segrega
40	99.5	5.9	1165	2 A97457	structural maintena
41	98	5.8	869	2 H83500	probable glucosyl
42	97.5	5.7	479	2 S75693	hypothetical prote
43	97	5.7	1774	2 T17421	polyketide synthas
44	96.5	5.7	914	2 S18942	hypothetical prote
45	95	5.6	963	2 S45167	chitin synthase (E

ALIGNMENTS

RESULT 1

JC4812

hyaluronan synthase (EC 2.4.1.1.-) - human

C:Species: Homo sapiens (man)

C>Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 02-Mar-2001

C:Accession: JC4812

R:Itano, N.; Kimata, K.

Biochem. Biophys. Res. Commun. 222, 816-820, 1996

A:Title: Molecular cloning of human hyaluronan synthase.

A:Reference number: JC4812; MUID:96244584; PMID:8651928

A:Accession: JC4812

A:Molecule type: mRNA

A:Residues: 1-543 <ITA>

A:Cross-references: DBEJ:D84424; NID:G1401033; PIDN:BAAL2351.1; PID:d1013030; PID:g1401034

C:Comment: This enzyme synthesizes a high molecular weight glycosaminoglycan composed of

C:Genetics:

A:Gene: GDB:HAS1; HAS

A:Cross-references: GDB:1220109; OMIM:601463

A:Map position: 19q13.4-19q13.4

C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein

F:58,167,308,321/Binding site: phosphate (Thr) (covalent) #status predicted

F:82,247/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 91.4%; Score 1550.5; DB 2; Length 543;

Best Local Similarity 92.8%; Pred. No. 7.6e-128;

Matches 297; Conservative 3; Mismatches 1; Indels 19; Gaps 1;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA 60

Db 26 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA 85

QY 61 LLYPRARVRLVMVDGNRAEDLYMVDMPREVFADDDPATYVWDGNYHQFWEPAAGAVGA 120

Db 86 LLYPRARVRLVMVDGNRAEDLYMVDMPREVFADDDPATYVWDGNYHQFWEPAAGAVGA 145

QY 121 GAYREVEADPGRLAVALVTRRCVCVAQRWGGRVYMTAFKALGDSVDVQVCDSDT 180

Db 146 GAYREVEADPGRLAVALVTRRCVCVAQRWGGRVYMTAFKALGDSVDVQVCDSDT 205

QY 181 RLDPMLLELVRVLDDEPRVGAVGDDVRILNPLDSWSVFLSSRLRYWVAFNVERACQSYFH 240

Db 206 RLDPMLLELVRVLDDEPRVGAVGDDVRILNPLDSWSVFLSSRLRYWVAFNVERACQSYFH 265

QY 241 CVCSISGPLGLYRNLLQOFLFAMYNQKFLGTHCTFGDDRHLTNRMLSGVATK----- 294

Db 266 CVCSISGPLGLYRNLLQOFLFAMYNQKFLGTHCTFGDDRHLTNRMLSGVATKTSRSR 325

QY 295 -----AEGTWS 301

Db 326 CYSTPSSFLRWLSQQTWS 345

```
RESULT 2
A43740
C:Species: Xenopus laevis (African clawed frog)
C>Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 09-Jul-2004
C:Accession: A43740
R:Roza, F.; Sargent, T.D.; Rebbert, M.L.; Michaelis, G.S.; Jamrich, M.; Grunz, H.; Jonas,
Dev. Biol. 129, 114-123, 1988
A>Title: Accumulation and decay of DG42 gene products follow a gradient pattern during X
A:Reference number: A43740; MUID:88313363; PMID:3410156
A:Accession: A43740
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-588 <ROS>
A:Cross-references: UNIPROT:P13563; EMBL:M222449; NID:g214098; PID:AAA49699.1; PID:g2140

Query Match 56.9%; Score 965.5; DB 2; Length 588;
Best Local Similarity 60.5%; Pred. No. 1.4e-76;
Matches 184; Conservative 47; Mismatches 62; Indels 11; Gaps 4;

QY 6 HLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLROCLASARALLYPR 65
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 HLMQSLFAFLAIRV---NKSELPCKFKTVALTIAGYQENPEYLIRKLESCKTVKYPK 128
QY 66 ARLRLVMVVDGNRAEDLYMVDMFREVFADDPATVYWDGNYHQWPWEPAAAGAVGAGAYRE 125
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 DKLKILVIDGNTEDDATMMEMFKDVFHGEDVGTIVWKGNYHTVVKPE---ETNKGSCPE 185
QY 126 VEA---EDPGLRLAVALVTRRCVCVAQRWGGKREVMYTFKALGDSVDYVQVCDSDTRL 182
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 VSKPLNEDEGINVVEELVRNKEVCVMQMGKGRVMTAFQAIGTSVDYVQVCDSDTKL 245
QY 183 DPMALLELVRLDDEPRVGAAGVDVRIINPLDSWVSFLSSRLYVWAFVNERACQSYFHCV 242
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
246 DELATVEMVKVLESNDMTYGAAGVDVRIINPNYDSFISFMSLSRLYVWAFVNERACQSYFDCV 305
QY 243 SCISGLPLGYRNLLQOFLAWYNQKFLGTHCTFGDDRHLTNRLMSMGVATK--AEGTRW 300
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 SCISGLPLGYRNLLQVFLAWYRQKFLGTYCTLGGDRHLTNRLMSMGRTYKTHKRAF 365
QY 301 SGP 304
Db |||
366 SGP 369

RESULT 3
TI7588
hYaluronoglucosaminidase-like protein - Chlorella virus PBCV-1
N:Alternate names: hYalurononic acid synthetase
C:Species: Chlorella virus PBCV-1
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: TI7588
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: TI7588
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-568 <GRA>
A:Cross-references: UNIPROT:Q84419; EMBL:U42580; NID:g4028896; PIDN:AAC96466.1
A:Note: A98R

Query Match 24.1%; Score 409; DB 2; Length 568;
Best Local Similarity 32.9%; Pred. No. 7.4e-28;
Matches 103; Conservative 63; Mismatches 105; Indels 42; Gaps 10;

QY 2 FLSAHLVAQSLFAYLEHRR----VAAARGPLDAATARSVALTISAYQEDPAYLROCLAS 57
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 VFVGFLLAQVLFSELNRKRLRWISLRPKGWINDV----RLAVIIAGYREDPYMFKLCLES 111
QY 58 ARALLYPRARLVLMVVDGNRAEDLYMVDMFREVFADDPATVYWDGNYHQWPWEPAAAGA 117
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 112 VRSDYGNV-ARLLICVIDGDEDDDMRMAAVYKAIYND-----NIKKP----- 152
QY 118 VGAGAYREVEAEDPGRRLAVEALVTRRCVCVAQRWGGKREVMYTFK--ALGDSVDYVQV 175
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 ----EFVLCESDDKEGERIDS--DFSRDICVLQPHRGKRECLYTGFLAKMDPPSVNAVVL 206
QY 176 CDSTRLDPMALLELVRLDDEPRVGAAGVDVRIINPLDSWVSFLSSRLYVWAFVNERAC 235
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
207 IDSTVLEKDAILEVVYPLACDPFIQAVAGECKLWN-TDILLSLLVAWRYYSACVERSA 265
QY 236 QSYFHCVCSCISGLGYRNLLQOFLAWYNQKFLGTHCTFGDDRHLTNRLMSMG-----Y 291
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 QSFRTVQCVGGPLGAYKIDIIKEIKPWISQRFQKCTYGGDRRLTNEILMRGKKVVF 325
QY 292 ATKAEGRWSTGP 304
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 TPFVAVG--WSDSP 336

RESULT 4
ZZZRCL
modulation protein nodC - Rhizobium leguminosarum plasmid pRL1J1
C:Species: Rhizobium leguminosarum
C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A03486
R:Rozen, L.; Johnston, A.W.B.; Downie, J.A.
Nucleic Acids Res. 12, 9497-9508, 1984
A>Title: DNA sequence of the Rhizobium leguminosarum nodulation genes nodAB and C requir
A:Reference number: A03482; MUID:85087952; PMID:6514582
A:Accession: A03486
A:Molecule type: DNA
A:Residues: 1-424 <ROS>
A:Cross-references: UNIPROT:P04340; GB:X01650; NID:g46212; PIDN:CAA8619.1; PID:g46215
C:Comment: This is one of the proteins, coded by nodulation genes, that are required for
C:Genetics:
A:Gene: nodC
A:Genome: plasmid
C:Superfamily: nodulation protein nodC
C:Keywords: nodulation

Query Match 17.2%; Score 292; DB 1; Length 424;
Best Local Similarity 26.6%; Pred. No. 9.1e-18;
Matches 84; Conservative 47; Mismatches 113; Indels 72; Gaps 8;

QY 7 LVAQSLFAYLE-----HRRVAAAARGP---LDAATARSVALTISAYQEDPAYLRQC 54
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9 IAAISLYAMLSVTVYKSAQVFHARETTISTTPAKDIETNPVPSVDVIVPCFNEDEPIVLSEC 68
QY 55 LASARALLYPRARLVLMVVDGNRAEDLYMVDMFREVFADDPATVYWDGNYHQWPWEPAA 114
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 LASLAEQDY-AGKURIVYVDGSKNRDAVVAQ--RAAYADDERENFT----- 112
QY 115 AGAVGAGAYREVEAEDPGRRLAVEALVTRRCVCVAQRWGGKREVMYTFKALGDSVDYVQ 174
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 -----ILPKNVGKRKAIAAITQSSGDLILNV- 138
QY 175 VCDSDTRLDPMALLELVRLDDEPRVGAAGVDVRIINPLDSWVSFLSSRLYVWAFVNERA 234
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 --DSDTTIAPDVWSKLAKHM-RDPAVGAAMQMKASNAQADTWLTRLIDMEYWLACNEERA 195
QY 235 QSYFHCVCSCISGLGYRNLLQOFLAWYNQKFLGTHCTFGDDRHLTNRLMSMGVATK 294
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 AQARFGAVMCCGPGCAMYRRSAMLSDIDQYETQLYRGKPSDFGEDRHLTILMSLGGFTE 255
QY 295 AEGTRWSTGTPPGAAT 310
Db |||
256 Y-----VPSAIAAT 264

RESULT 5
A95321
NodC N-ACETYLGUCOSAMINYLTRANSFERASE [imported] - Sinorhizobium meliloti (strain 1021) m
C:Species: Sinorhizobium meliloti
```


Db 3 VAVVPSYNESASAIWNTINSVLAQDYPIH--EIPFVDDGSKDKGAYEVALKMRBELRT 60

QY 95 EDPATYVWDGNYHQWPAPAAAGAGAGAYREVEAEDPGRGLAVEALVTRRCVC----- 147

Db 61 Q-----REIAA-----TTKNICSEILGIP 79

QY 148 --VAQRWG---KKREVMYTAFKALGSDVDYVQVCDSDTRLDPMALLELVRLVDEDPVGA 202

Db 80 DLIVHRLPKNCGRKHAQLWAFR--TTDAIVTIDSDGLFFNAVRELLKPFN-DEKVA 136

QY 203 VGGDVRILNPLDSWSFLSSRLYVAFNVERACQSFVHCVSCISGPLGLYRNLLAQOFL 262

Db 137 TTGHVNRNRNDNLITKLIDMDYDIAFRVERAAQSVTGNLVCSGPLSCYRREVITENLE 196

QY 263 AWYQKFLGTHCTFGDDRHLTNRLMSG 290

Db 197 HYGSMFLGEEVQFGDDRCCLTNVAILKG 224

RESULT 11

A48755

hyaluronan synthase (EC 2.4.1.-) - Streptococcus pyogenes

C:Species: Streptococcus pyogenes

C>Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 09-Jul-2004

C:Accession: A48755; JC2077

R:DeAngelis, P.L.; Papaconstantinou, J.; Weigel, P.H.

J. Biol. Chem. 268, 19181-19184, 1993

A:Title: Molecular cloning, identification, and sequence of the hyaluronan synthase gene

A:Reference number: A48755; MUID:93374890; PMID:8366070

A:Accession: A48755

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-395 <DEA1>

A:Cross-references: UNIPROT:Q54866; UNIPROT:Q54867; UNIPROT:Q8K5G6; GB:L20853; NID:94100

R:De Angelis, P.L.; Yang, N.; Weigel, P.H.

Biochem. Biophys. Res. Commun. 199, 1-10, 1994

A:Title: The Streptococcus pyogenes hyaluronan synthase: Sequence comparison and conservation

A:Reference number: JC2077; MUID:94168559; PMID:8122999

A:Accession: JC2077

A:Molecule type: DNA

A:Residues: 1-395 <DEA2>

C:Comment: This enzyme is responsible for polymerizing an alternating polysaccharide copolymer

C:Genetics:

A:Gene: hasA

C:Superfamily: modulation protein nodC

C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

F:9-29/Domain: transmembrane #status predicted <TM1>

F:294-314/Domain: transmembrane #status predicted <TM2>

F:328-341/Domain: transmembrane #status predicted <TM3>

F:353-370/Domain: transmembrane #status predicted <TM4>

Query Match 12.6%; Score 213.5; DB 2; Length 395;

Best Local Similarity 28.1%; Pred. No. 6.3e-11;

Matches 74; Conservative 27; Mismatches 95; Indels 67; Gaps 6;

QY 37 VALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAEDLYM-----VDM 87

Db 42 VAAVPSYNEDEASLETLKSLVLAQTYPLS--EIIIVDDGSSNTDAIQIIEEYVNVREVDI 99

QY 88 FREVPADDDPATYVWDGNYHQWPAPAAAGAGAGAYREVEADPGRGLAVEALVTRRCVC 147

Db 100 CRNVIVHR---SLVNKGRKH----- 116

QY 148 VAQRWGGKREVMYTAFAKLGSDVDYVQVCDSDTRLDPMALLELVRLVDEDPVGA 207

Db 117 -AQAWAFER-----SDADVLTVDSDTYIYPNALELLKSFN-DETVYAATGHL 163

QY 208 RILNPLDSWSFLSSRLYVAFNVERACQSFVHCVSCISGPLGLYRNLLAQOFL 267

Db 164 NARNRQTNLLTRLTDIRVDNAGVERAAQSLTGNILVCSGPLSIYRREVIIPNLERYKNQ 223

QY 268 KFLGTHCTFGDDRHLTNRLMSG 290

Db 224 TFLGLPVSIGDDRCCLTNVAILDG 246

RESULT 12

A53100

hyaluronate synthase A (HasA) - Streptococcus sp. (group A)

C:Species: Streptococcus sp.

C>Date: 06-Oct-1994 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C:Accession: A53100

R:Dougherty, B.A.; van de Rijn, I.

J. Biol. Chem. 269, 169-175, 1994

A:Title: Molecular characterization of hasA from an operon required for hyaluronic acid synthesis

A:Reference number: A53100; MUID:94103204; PMID:8276791

A:Accession: A53100

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-419 <DOU>

A:Cross-references: UNIPROT:Q8NKK1

A>Note: sequence extracted from NCBI backbone (NCBIN:141683, NCBIP:141684)

C:Superfamily: modulation protein nodC

Query Match 12.6%; Score 213.5; DB 2; Length 419;

Best Local Similarity 28.1%; Pred. No. 6.7e-11;

Matches 74; Conservative 27; Mismatches 95; Indels 67; Gaps 6;

QY 37 VALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAEDLYM-----VDM 87

Db 66 VAAVPSYNEDEASLETLKSLVLAQTYPLS--EIIIVDDGSSNTDAIQIIEEYVNVREVDI 123

QY 88 FREVPADDDPATYVWDGNYHQWPAPAAAGAGAGAYREVEADPGRGLAVEALVTRRCVC 147

Db 124 CRNVIVHR---SLVNKGRKH----- 140

QY 148 VAQRWGGKREVMYTAFAKLGSDVDYVQVCDSDTRLDPMALLELVRLVDEDPVGA 207

Db 141 -AQAWAFER-----SDADVLTVDSDTYIYPNALELLKSFN-DETVYAATGHL 187

QY 208 RILNPLDSWSFLSSRLYVAFNVERACQSFVHCVSCISGPLGLYRNLLAQOFL 267

Db 188 NARNRQTNLLTRLTDIRVDNAGVERAAQSLTGNILVCSGPLSIYRREVIIPNLERYKNQ 247

QY 268 KFLGTHCTFGDDRHLTNRLMSG 290

Db 248 TFLGLPVSIGDDRCCLTNVAILDG 270

RESULT 13

JC6079

chitin synthase (EC 2.4.1.16) chsD - Emericella nidulans

C:Species: Emericella nidulans, Aspergillus nidulans

C>Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 20-Jun-2000

C:Accession: JC6079

R:Mocoyama, T.; Fujiwara, M.; Kojima, N.; Horiuchi, H.; Ohta, A.; Takagi, M.

Mol. Gen. Genet. 251, 442-450, 1996

A:Title: The Aspergillus nidulans genes chsA and chsD encode chitin synthases which have

A:Reference number: JC6079; MUID:96285568; PMID:8709948

A:Accession: JC6079

A:Molecule type: DNA

A:Residues: 1-1086 <MOT>

A:Cross-references: DDBJ:D83246; NID:gl688025; PIDN:BAAL1866.1; PID:gl688026

C:Comment: This enzyme functions in conidia formation.

C:Genetics:

A:Gene: chsD

A:Introns: 800/2

C:Superfamily: chitin synthase chs4

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 10.9%; Score 185; DB 2; Length 1086;

Best Local Similarity 22.0%; Pred. No. 6.7e-08;

Matches 71; Conservative 54; Mismatches 131; Indels 56; Gaps 11;

QY 28 PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDG---NRAEDLYM 84

Db 563 PYGPPLAHAMCL-VTCYSEGEIGRTTLDLSIALTDYPNSHKSIIVVICDGIILKNGEBFST 621
Qy 85 VDMFREVAD-----ED-----PATYVDGNVHQHPWEPAAAGAVGAGAYREVEADPGRLA 135
Db 622 PDIVLMMRDPIIPPEEYEAFFYAVATGSKRHNNMAKYAGFYDGEHSIIPVENMQQRVP 681
Qy 136 VEALVRRRCVCVAQRWG-----GKR-----EVMYTAFKAL--- 166
Db 682 MMIVV---KCGTPAARTAAKPNRGKRDSQIILMSFLQKVPFDERMTELEYEMFNGLLHV 738
Qy 167 ----GDSVDYVQVCDSTRDPMALLELVRLVLDPRVGAGVDVRLNPLDLSWVSFLSS 222
Db 739 TGIPDPDFEVLVMDADTKVPFDSLTHMSAMVKDPEVMGLCGETKIANKTDSWVTMIQV 798
Qy 223 LRYWAFNVERACQSYFHCVCISGSLGLYRNLLQOFLAW-----YNQKFL 270
Db 799 FEYFVSHHQKAFESVFGVTCFPCFSMYRIKAPKGGQNTWVPILANPDVIVHSENVV 858
Qy 271 GT-H----CTFGDDRHLNRM 287
Db 859 DTLHKNNLLLGEDRYLSTLML 880

RESULT 14

S61896

chitin synthase (EC 2.4.1.16) CHS4 - Neurospora crassa

N:Alternate names: chitin synthase class IV

C:Species: Neurospora crassa

C:Date: 23-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: S61886

R:Din, A.B.; Specht, C.A.; Robbins, P.W.; Yarden, O.

Mol. Gen. Genet. 250, 214-222, 1996

A:Title: chs-4, a class IV chitin synthase gene from Neurospora crassa.

A:Reference number: S61886; MUID:96188842; PMID:8628221

A:Accession: S61886

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-1195 <DIN>

A:Cross-references: UNIPROT:Q01285; EMBL:U25097; NID:G793937; PIDN:AAB03563.1; PID:G7939

C:Genetics:

A:Gene: chs-4

A:Introns: 1042/1

C:Superfamily: chitin synthase chs4

C:Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase; transmembrane

Query Match 10.4%; Score 177; DB 2; Length 1195;

Best Local Similarity 22.4%; Pred. No. 3.8e-07;

Matches 72; Conservative 54; Mismatches 130; Indels 66; Gaps 12;

Qy 28 PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDG---NRAEDLYM 84
Db 641 PFGPPLAHTICL-VTAYSEGBMGVRTLDLSIAMDYFNHSHKVLIVICDGIILKNGHEHST 699
Qy 85 VDMFREVAD-----DPATY--VMDGNVHQHPWEPAAAGAVGAGAYREVEADPGRLA 135
Db 700 PDILGLMKDKHTIHPDDVEPFSYAVATGSKRHNNMAKYTYGTYGNTSAIPLKQQRVP 759
Qy 136 VEALVRRRCVCVAQRW-----GKR-----EVMYTAFKAL--- 166
Db 760 MMVVV---KCGTPAEASKPKGNRGKRDSQIILMSFLQKVPFDERMTELEYEMFNGLLWKI 816
Qy 167 -GDSVDYVQV---CDSTRDPMALLELVRLVLDPRVGAGVDVRLNPLDLSWVSFLSS 222
Db 817 TGISPDFFEYVLMVDADTKVPFDSLTHMSAMVKDPEIMGLCGETKIANKRASWVSAIQV 876
Qy 223 LRYWAFNVERACQSYFHCVCISGSLGLYRNLLQOFLAW-----YNQKFL 270
Db 877 FEYFVSHHLAKAFESVFGVTCFPCFCMYRIKAPKGGQNTWVPILANPDVIVHSENVV 936
Qy 271 GT-H----CTFGDDRHLNRM 287
Db 937 DTLHKNNLLLGEDRYLSTLML 958

RESULT 15

S45879

chitin synthase (EC 2.4.1.16) 3 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YBR023c; protein YBR0305

C:Species: Saccharomyces cerevisiae

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: S45879; S46554; S22776; A39639; S17247

R:Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45875

A:Accession: S45879

A:Molecule type: DNA

A:Residues: 1-1165 <GRI>

A:Cross-references: UNIPROT:P29465; EMBL:Z35892; NID:G536229; PIDN:CAA84965.1; PID:G5362

A:Experimental source: strain S288C

R:Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.

Yeast 10(Suppl.A), S75-S80, 1994

A:Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II from

ly identified genes and a homologue of the SCO1 gene.

A:Reference number: S46551; MUID:94378725; PMID:8091864

A:Accession: S46554

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1165 <SMI>

A:Cross-references: EMBL:X76078; NID:G498748; PIDN:CAAS3680.1; PID:G498752

A:Experimental source: strain S288C

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993

R:Bulawa, C.E.

Mol. Cell. Biol. 12, 1764-1776, 1992

A:Title: CSD2, CSD3, and CSD4, genes required for chitin synthesis in Saccharomyces cere

bium species and Xenopus laevis.

A:Reference number: S22776; MUID:92195323; PMID:1532231

A:Accession: S22776

A:Molecule type: DNA

A:Residues: 1-1162,'L',1164-1165 <BUL>

A:Cross-references: EMBL:M73697; NID:G172103; PIDN:AAA34844.1; PID:G172104

R:Valdivieso, M.H.; Mol, P.C.; Shaw, J.A.; Cabib, E.; Duran, A.

J. Cell Biol. 114, 101-109, 1991

A:Title: CAL1, a gene required for activity of chitin synthase 3 in Saccharomyces cerevi

A:Reference number: A39639; MUID:91268144; PMID:2050737

A:Accession: A39639

A:Molecule type: DNA

A:Residues: 67-1165 <VAL>

A:Cross-references: GB:X57300; NID:G3359; PIDN:CAA40559.1; PID:G3360

C:Genetics:

A:Gene: SGD:CHS3; CAL1; CSD2; MIPS:YBR023c

A:Cross-references: MIPS:YBR023c; SGD:S0000227

A:Map position: 2R

C:Function:

A:Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosa

C:Superfamily: chitin synthase chs4

C:Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase; transmembrane

F:203-219/Domain: transmembrane #status predicted <TM1>

F:457-473/Domain: transmembrane #status predicted <TM2>

F:1018-1034/Domain: transmembrane #status predicted <TM3>

F:1035-1054/Domain: transmembrane #status predicted <TM4>

F:1060-1076/Domain: transmembrane #status predicted <TM5>

F:1084-1100/Domain: transmembrane #status predicted <TM6>

Query Match 10.3%; Score 174; DB 1; Length 1165;

Best Local Similarity 21.1%; Pred. No. 6.7e-07;

Matches 67; Conservative 59; Mismatches 108; Indels 84; Gaps 10;

Qy 41 ISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDG---NRAED-----LYMVDMPRE 90
Db 658 VTCYSEDEGLRTTLDLSLTDDYFNHSHKLMVVDGLIKSGNDKTTPEIAGMWDDEFT 717
Qy 91 VFADEDPATYVDGNVHQHPWEPAAAGA-----VGAGAYR-----EVEADPGRLAVEAL 139
Db 718 PPDEVKPSYV-----AVASGSKRHNNMAKYAGFYKDDSTIPPENQQRVPIT 767
Qy 140 VRTRRCVCAQRWG-----GKR--EVMYTAFA-----KAL 166

Db 768 V---KCGTPAEQGAAPGNRGRDSQIILMSFLEKITFDERMTQLEFOLLKNIWQITGLM 824
Qy 167 GDSVDYVOVCDSDTRLDPMALLELVRLVLEDDPRVGAVGDDVRILNPLDSWVSFLSSLRW 226
Db 825 ADPIETVLMDADTKVPDPALTHMVAEMVKDPLIMGLCGETKIANKAQSWTQVFEYY 884
Qy 227 VAFNVERACQSYFHCVCISGELGLYR-----NNLLQQFLEAWYNQKF 269
Db 885 ISHQAKAFESVFGSVTCLPGCFSMYRIKSPKGSQYWPVLANPDIVERYSNDVNTLH 944
Qy 270 LGTHCTFGDDRHLTNRML 287
Db 945 KKNLLLGEDRFLSSIML 962

Search completed: March 11, 2005, 14:23:14
Job time : 19.7879 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:48:17 ; Search time 83.6364 Seconds
(without alignments)
1959.259 Million cell updates/sec

Title: US-10-672-399-8

Perfect score: 1697

Sequence: 1 AFLSAHLVAQSLFAYLEHRR.....SGTPGPAAQRRPRPSCGG 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1553.5	91.5	578	HAS1_HUMAN	Q92839 homo sapien
2	1553.5	91.5	578	Q9NS49	Q9ns49 homo sapien
3	1550.5	91.4	577	Q81YH3	Q81yh3 homo sapien
4	1508.5	88.9	583	HAS1_MOUSE	Q61647 mus musculus
5	1508	88.9	583	Q6S742	Q6s742 papio anubi
6	1505.5	88.7	583	Q8CH93	Q8ch93 rattus norv
7	1048.5	61.8	458	Q6T488	Q6t488 brachydanio
8	965.5	55.9	588	HAS1_XENLA	P13563 xenopus lae
9	948.5	55.9	582	Q95M29	Q95m29 oryctolagus
10	938	55.3	553	Q75R37	Q75r37 sus scrofa
11	937	55.2	552	Q8S070	Q8s070 sus scrofa
12	935.5	55.1	554	HAS3_MOUSE	Q8650 mus musculus
13	935	55.1	552	HAS2_MOUSE	P70312 mus. musculus
14	935	55.1	552	Q8HZJ3	Q8hzj3 equus caball
15	935	55.1	552	Q95M16	Q95m16 oryctolagus
16	934	55.0	552	HAS2_HUMAN	Q92819 homo sapien
17	933.5	55.0	554	Q8CEB9	Q8ceb9 mus musculus
18	933.5	55.0	554	Q8CH92	Q8ch92 rattus norv
19	933	55.0	552	HAS2_BOVIN	Q97711 bos taurus
20	930	54.8	552	HAS2_RAT	Q35776 rattus norv
21	928.5	54.7	554	Q3DG40	Q3dg40 brachydanio
22	928	54.7	553	HAS3_HUMAN	Q00219 homo sapien
23	928	54.7	553	Q36R72	Q36r72 homo sapien
24	927	54.6	557	Q8W9J2	Q8w9j2 xenopus lae
25	924	54.4	552	Q3DG41	Q3dg41 brachydanio
26	923	54.4	552	HAS2_CHICK	Q57424 gallus galli
27	907	53.4	551	HAS2_XENLA	O57427 xenopus lae
28	843.5	49.7	583	HAS3_XENLA	O57428 xenopus lae
29	842.5	49.6	583	Q8AZI0	Q8azi0 xenopus lae
30	742.5	43.8	393	Q811Y6	Q81ly6 rattus norv
31	552.5	32.6	156	O18792	O18792 papio anubi

32 528.5 31.1 281 2 Q8WTZ0 Q8wtz0 homo sapien
33 488 28.8 134 2 Q90489 Q90489 brachydanio
34 473.5 27.9 245 2 Q9GK14 Q9gk14 bos taurus
35 409 24.1 568 2 Q84119 Q84119 paramesicium
36 409 24.1 568 2 Q9WFS9 Q9wfs9 paramesicium
37 409 24.1 568 2 Q9WFT0 Q9wft0 paramesicium
38 355.5 20.9 172 2 Q9XS25 Q9xs25 bos taurus
39 301.5 17.8 131 2 Q8BPN0 Q8bpn0 mus musculus
40 292 17.2 424 1 NODC_RHILV P04340 rhizobium 1
41 290.5 17.1 452 2 Q9AQZ3 Q9aqz3 bradyrhizob
42 279.5 16.5 443 2 Q8KLG3 Q8klg3 rhizobium e
43 279 16.4 426 1 NODC_RHIME P04341 rhizobium m
44 276.5 16.3 452 2 Q6EX51 Q6ex51 sinorhizobi
45 274 16.1 402 2 O52478 O52478 rhizobium m

ALIGNMENTS

RESULT 1

ID HAS1_HUMAN STANDARD; PRT; 578 AA.
AC Q92839; Q14470;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
GN Name=HAS1; Synonyms=HAS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RX MEDLINE=96394438; PubMed=8798544; DOI=10.1074/jbc.271.38.23395;
RA Shyjan A.M., Heidlin P., Butcher E.C., Yoshino T., Brisken M.J.;
RT "Functional cloning of the cDNA for a human hyaluronan synthase.";
RL J. Biol. Chem. 271:23395-23399(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96244584; PubMed=8651928; DOI=10.1006/bbrc.1996.0827;
RA Itano N., Kinata K.;
RT "Molecular cloning of human hyaluronan synthase.";
RL Biochem. Biophys. Res. Commun. 222:816-820(1996).
CC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA) synthesis. Also able to catalyze the synthesis of chito-oligosaccharide depending on the substrate.
CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-glucuronosyl(1->3)](n) + 2n UDP.
CC -!- COFACTOR: Magnesium.
CC -!- PATHWAY: Hyaluronate synthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: Highly expressed in ovary followed by spleen, thymus, prostate, testes and large intestine. Weakly expressed in small intestine.
CC -!- SIMILARITY: Belongs to the nodC/HAS family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U59269; AAC50706.1; -;
EMBL; D84424; BAA12351.1; ALT_INIT.
Genew; HGNC:4818; HAS1.
DR MIN; 601463; -.

DR GO: 0005887; C: integral to plasma membrane; TAS.
 DR GO: 0007155; P: cell adhesion; TAS.
 DR GO: 0006024; P: glycosaminoglycan biosynthesis; TAS.
 DR InterPro; IPR001173; Glyco_trans_2.
 DR Pfam; PF09535; Glycosyltransferase; Multigene family; Transferase; Transmembrane.
 KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
 FT DOMAIN 1 25 Cytoplasmic (Potential).
 FT TRANSMEM 26 46 1 (Potential).
 FT DOMAIN 47 52 Extracellular (Potential).
 FT TRANSMEM 53 73 2 (Potential).
 FT DOMAIN 74 399 Cytoplasmic (Potential).
 FT TRANSMEM 400 420 3 (Potential).
 FT DOMAIN 421 430 Extracellular (Potential).
 FT TRANSMEM 431 451 4 (Potential).
 FT DOMAIN 452 457 Cytoplasmic (Potential).
 FT TRANSMEM 458 478 5 (Potential).
 FT DOMAIN 479 497 Extracellular (Potential).
 FT TRANSMEM 498 518 6 (Potential).
 FT DOMAIN 519 540 Cytoplasmic (Potential).
 FT TRANSMEM 541 561 7 (Potential).
 FT DOMAIN 562 578 Extracellular (Potential).
 FT CONFLICT 1 2 MR -> RS (in Ref. 2).
 FT CONFLICT 34 34 G -> A (in Ref. 2).
 SQ SEQUENCE 578 AA; 64884 MW; 355FD54B0899E43C CRC64;

Query Match 91.5%; Score 1553.5; DB 1; Length 578;
 Best Local Similarity 93.1%; Pred. No. 6.2e-125;
 Matches 298; Conservative 2; Mismatches 1; Indels 19; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
 Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
 Qy 61 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQWPEPAAAGAVGA 120
 Db 121 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQWPEPAAAGAVGA 180
 Qy 121 GAYREVEADPGRLAVEALVTRRCVCVAQWGWGKREVMYAFKALGSDYVQVCDSDT 180
 Db 181 GAYREVEADPGRLAVEALVTRRCVCVAQWGWGKREVMYAFKALGSDYVQVCDSDT 240
 Qy 181 RLDPMALLEVRVLDDEDPVGAAGDVRILNPLDSWVSLSSLYWVAFNVERACQSYFH 240
 Db 241 RLDPMALLEVRVLDDEDPVGAAGDVRILNPLDSWVSLSSLYWVAFNVERACQSYFH 300
 Qy 241 CVSCISGPLGLYRNLLQQFLYVWYVQVCDSDT 294
 Db 301 CVSCISGPLGLYRNLLQQFLYVWYVQVCDSDT 360
 Qy 295 -----AEGTRWS 301
 Db 361 CYSETPSSFLRWLSQOTRWS 380

RESULT 2
 Q9NS49 ID Q9NS49 PRELIMINARY; PRT; 578 AA.
 AC Q9NS49;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hyaluronan synthase 1.
 GN Name=HAS1;
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
 RA Phan H., Velasco N., Do L., Regalla W., Terry A., Brower A., Garnes J.,
 RA Danganan L., Erlar A., Christensen M., Georgescu A., Avila J., Liu S.,

RA Andreise T., Trankheim M., Attix C., Amico-Keller G., Coefield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carrano A.V.,
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA DOE Joint Genome Institute;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E.,
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC018755; AA87845.1; -;
 DR InterPro; IPR001173; Glyco_trans_2.
 DR InterPro; IPR002076; GPCR_Rhodopsin.
 DR InterPro; IPR002057; Isopen N synth.
 DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; UNKNOWN_1.
 DR PROSITE; PS00185; IPNS 1; UNKNOWN_1.
 SQ SEQUENCE 578 AA; 64831 MW; 2FE3A44B0D5380FF CRC64;

Query Match 91.5%; Score 1553.5; DB 2; Length 578;
 Best Local Similarity 93.1%; Pred. No. 6.2e-125;
 Matches 298; Conservative 2; Mismatches 1; Indels 19; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
 Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
 Qy 61 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQWPEPAAAGAVGA 120
 Db 121 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQWPEPAAAGAVGA 180
 Qy 121 GAYREVEADPGRLAVEALVTRRCVCVAQWGWGKREVMYAFKALGSDYVQVCDSDT 180
 Db 181 GAYREVEADPGRLAVEALVTRRCVCVAQWGWGKREVMYAFKALGSDYVQVCDSDT 240
 Qy 181 RLDPMALLEVRVLDDEDPVGAAGDVRILNPLDSWVSLSSLYWVAFNVERACQSYFH 240
 Db 241 RLDPMALLEVRVLDDEDPVGAAGDVRILNPLDSWVSLSSLYWVAFNVERACQSYFH 300
 Qy 241 CVSCISGPLGLYRNLLQQFLYVWYVQVCDSDT 294
 Db 301 CVSCISGPLGLYRNLLQQFLYVWYVQVCDSDT 360
 Qy 295 -----AEGTRWS 301
 Db 361 CYSETPSSFLRWLSQOTRWS 380

RESULT 3
 Q8IYH3 ID Q8IYH3 PRELIMINARY; PRT; 577 AA.
 AC Q8IYH3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hyaluronan synthase 1.
 GN Name=HAS1;
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Ovary;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywicki M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
SEQUENCE FROM N.A.
TISSUE=Ovary;
R Strausberg R.;
Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
EMBL; BC035837; AAH35837.1 ; -
InterPro; IPR001173; Glyco trans. 2.
InterPro; IPR00276; GPCR_Rhodospn.
InterPro; IPR02057; Isopen N synth.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_1; UNKNOWN_1.
DR PROSITE; PS00185; IPNS_1; UNKNOWN_1.
SQ SEQUENCE 577 AA; 64788 MW; 61C9AC33AEBD183 CRC64;
Query Match 91.4%; Score 1550.5; DB 2; Length 577;
Beat Local Similarity 92.8%; Pred. No. 1.1e-124;
Matches 297; Conservative 3; Mismatches 1; Indels 19; Gaps 19
OY 1 AFLSAHLVAQLSFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCCLASARA 60
Db 60 AFLSAHLVAQLSFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCCLASARA 119
OY 61 LLYPRLRLVLVMVDGNRAEDLYWDMFRVFPADEDPATYWDGNYHPWEAAPAAAGAVGA 120
Db 120 LLYPRLRLVLVMVDGNRAEDLYWDMFRVFPADEDPATYWDGNYHPWEAAPAAAGAVGA 179
OY 121 GAYREVEAEDPGLRALVEALVRTRRCVCVAQRWGKGKRVMTAFKALGDSDVDVVQCDSDT 180
Db 180 GAYREVEAEDPGLRALVEALVRTRRCVCVAQRWGKGKRVMTAFKALGDSDMDIVQCDSDT 239
OY 181 RLDPMALLEVRLVLEDPRVGAVGGDVRIILNPLDSWVSFLSLRYVFAFNVERACQSIFH 240
Db 240 RLDPMALLEVRLVLEDPRVGAVGGDVRIILNPLDSWVSFLSLRYVFAFNVERACQSIFH 299
OY 241 CVSCISGPLGYRNLLQQFLEAWYNOKFGTGHTCTGDDRHILTNRMLSMGYATK----- 294
Db 300 CVSCISGPLGYRNLLQQFLEAWYNOKFGTGHTCTGDDRHILTNRMLSMGYATKYTSRSR 359
OY 295 -----AEGTRWS 301 :::
Db 360 CYSTPSFRLWLSQLQTWS 379
RESULT 4
HASI MOUSE STANDARD; PRT; 583 AA.
ID HASI MOUSE Q61647;
AC Q61647;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hyaluronic acid synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE (Hyaluronic acid synthase 1) (HA synthase 1).
DG Name=Has1; Synonyms=Haa;
GN Mus musculus (Mouse);
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RNA [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=96215261; PubMed=8626618; DOI=10.1074/jbc.271.17.9875;

Itano N., Kimata K.;
"Expression cloning and molecular characterization of HAS protein, a
eukaryotic hyaluronan synthase.";
J. Biol. Chem. 271:9875-9878(1996).
[2]
MUTAGENESIS.
MEDLINE=20085071; PubMed=10617644; DOI=10.1074/jbc.275.1.497;
Yoshida M., Itano N., Yamada Y., Kimata K.;
"In vitro synthesis of hyaluronan by a single protein derived from
mouse HAS1 gene and characterization of amino acid residues essential
for the activity";
J. Biol. Chem. 275:497-506(2000).
-!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
synthesis. Also able to catalyze the synthesis of chito-
oligosaccharide depending on the substrate.
-!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-
glucuronosyl(1->3)](n) + 2n UDP.
-!- COFACTOR: Magnesium.
-!- PATHWAY: Hyaluronate synthesis.
-!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
-!- SIMILARITY: Belongs to the nOdc/HAS family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; D82964; BAA11654.1; -;
MGD; MGI:106590; Has1.
InterPro: IPR001173; Glyco trans 2.
DR PFam: PF00535; Glycos transf_2; I.
KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
DR DOMAIN 1 24 Cytoplasmic (Potential).
FT TRANSMEM 25 45 1 (Potential).
FT DOMAIN 46 51 Extracellular (Potential).
FT TRANSMEM 52 72 2 (Potential).
FT DOMAIN 73 404 Cytoplasmic (Potential).
FT TRANSMEM 405 425 3 (Potential).
FT DOMAIN 426 435 Extracellular (Potential).
FT TRANSMEM 436 456 4 (Potential).
FT DOMAIN 457 462 Cytoplasmic (Potential).
FT TRANSMEM 463 483 5 (Potential).
FT DOMAIN 484 501 Extracellular (Potential).
FT TRANSMEM 502 522 6 (Potential).
FT DOMAIN 523 545 Cytoplasmic (Potential).
FT TRANSMEM 546 566 7 (Potential).
FT DOMAIN 567 583 Extracellular (Potential).
FT TRANSMEM 81 84 Poly-Ala.
FT DOMAIN 519 Poly-Leu.
FT CARBOHYD 489 N-linked (GlcNAc...) (Potential).
FT MUTAGEN 242 D->E: Loss of both activities.
FT MUTAGEN 311 S->N: No effect.
FT MUTAGEN 312 G->P: No effect.
FT MUTAGEN 313 P->G: No effect.
FT MUTAGEN 314 L->V: Loss of HA activity.
FT MUTAGEN 314 L->I: 75% decrease of both activities.
FT MUTAGEN 344 D->E: Loss of both activities.
FT MUTAGEN 380 Q->N: 85%-90% decrease of both
activities.
FT MUTAGEN 383 R->K: 85%-90% decrease of both
activities.
FT MUTAGEN 384 W->Y: Loss of both activities.
FT SEQUENCE 583 AA; 65544 MW; 7AF9273E7B314728 CRC64;
Query Match 88.9%; Score 1508.5; DB 1; Length 583;
Best Local Similarity 89.9%; Pred. No. 4.7e-121;
Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALRTISAYQEDPAYLRQCL 55

```
Db 60 AFLSAHLVAQSLFAYLEHRRVAAAARRSLAKGLDAAATARSVALTISAYQEDPAYLRQCL 119
Qy 56 ASARALLYPRARLRVLMVVDGNRAEDLYMDMFREVFADDPATYVMDGNTHQWPEPA-A 114
Db 120 TSARALLYPHTRLRVLMVVDGNRAEDLYMDMFREVFADDPATYVMDGNTHQWPEPAEA 179
Qy 115 AGAVGAGAYREVEADPGRLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQ 174
Db 180 TGAVGEGAYREVEADPGRLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQ 239
Qy 175 VCDSDTRLDPMALLELVRLVLDDEPRVGAAGVDVRLINPLDSWVSLSLRYWVAFNVERA 234
Db 240 VCDSDTRLDPMALLELVRLVLDDEPRVGAAGVDVRLINPLDSWVSLSLRYWVAFNVERA 299
Qy 235 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 294
Db 300 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 359
Qy 295 -----AEGTRWS 301
Db 360 YTSRRCYSETPSSFLRWLSQOTRWS 385
```

RESULT 5

```
Q6S742 ID Q6S742 PRELIMINARY; PRT; 583 AA.
AC Q6S742;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hyaluronan synthase.
GN Name=HAS1;
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9555;
RN [1]
RP SEQUENCE FROM N.A.
RA Martinez-Duncker I., Oriol R., Mollicone R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY463695; AAR25554.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR002057; Isopen_N_synth.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00185; IPNS_1; UNKNOWN_1.
SQ SEQUENCE 583 AA; 65596 MW; EA47354C89346F94 CRC64;
```

Query Match 88.9%; Score 1508; DB 2; Length 583;
Best Local Similarity 89.5%; Pred. No. 5.1e-121;
Matches 291; Conservative 4; Mismatches 6; Indels 24; Gaps 2;

```
Qy 1 AFLSAHLVAQSLFAYLEHRRVA-----AARGPLDAAATARSVALTISAYQEDPAYLRQCL 55
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARRAAARGPLDAAATARSVALTISAYQEDPAYLRQCL 120
Qy 56 ASARALLYPRARLRVLMVVDGNRAEDLYMDMFREVFADDPATYVMDGNTHQWPEPAEA 115
Db 121 VSARALLYPRARLRVLMVVDGNRPEDLYMDMFREVFADDPATYVMDGNTHQWPEPAAV 180
Qy 116 GAVGAGAYREVEADPGRLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQ 175
Db 181 GAVGAGAYREVEADPGRLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQ 240
Qy 176 CDSSTRLDPMALLELVRLVLDDEPRVGAAGVDVRLINPLDSWVSLSLRYWVAFNVERAC 235
Db 241 CDSSTRLDPMALLELVRLVLDDEPRVGAAGVDVRLINPLDSWVSLSLRYWVAFNVERAC 300
Qy 236 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATK- 294
Db 301 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATKY 360
```

```
Qy 295 -----AEGTRWS 301
Db 361 TSRSRCYSETPSSFLRWLSQOTRWS 385

RESULT 6
Q8CH93 ID Q8CH93 PRELIMINARY; PRT; 583 AA.
AC Q8CH93;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hyaluronan synthase 1.
GN Name=HAS1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Itano N., Sawai T., Atsumi F., Miyaishi O., Taniguchi S., Kannagi R.,
RA Hamaguchi M., Kimata K.;
RT "Selective expression and functional characteristics of three
RT Mammalian hyaluronan synthases in oncogenic malignant
RT transformation.";
RL J. Biol. Chem. 279:18679-18687(2004).
DR EMBL; AB097568; BAC43730.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
DR InterPro; IPR002057; Isopen_N_synth.
DR PROSITE; PS00185; IPNS_1; UNKNOWN_1.
SQ SEQUENCE 583 AA; 65724 MW; 35513C6B21DE4B8D CRC64;
```

Query Match 88.7%; Score 1505.5; DB 2; Length 583;
Best Local Similarity 89.6%; Pred. No. 8.4e-121;
Matches 292; Conservative 2; Mismatches 7; Indels 25; Gaps 3;

```
Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAAATARSVALTISAYQEDPAYLRQCL 55
Db 60 AFLSAHLVAQSLFAYLEHRRVTVAAARRAFAGKPLDAAATARSVALTISAYQEDPAYLRQCL 119
Qy 56 ASARALLYPRARLRVLMVVDGNRAEDLYMDMFREVFADDPATYVMDGNTHQWPEPA-A 114
Db 120 TSARALLYPRTRLRVLMVVDGNRAEDLYMDMFREVFADDPATYVMDGNTHQWPEPAEA 179
Qy 115 AGAVGAGAYREVEADPGRLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQ 174
Db 180 AGAVGAGAYREVEADPGRLAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGDSVDYVQ 239
Qy 175 VCDSDTRLDPMALLELVRLVLDDEPRVGAAGVDVRLINPLDSWVSLSLRYWVAFNVERA 234
Db 240 VCDSDTRLDPMALLELVRLVLDDEPRVGAAGVDVRLINPLDSWVSLSLRYWVAFNVERA 299
Qy 235 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 294
Db 300 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 359
Qy 295 -----AEGTRWS 301
Db 360 YTSRRCYSETPSSFLRWLSQOTRWS 385
```

RESULT 7

```
Q6T488 ID Q6T488 PRELIMINARY; PRT; 458 AA.
AC Q6T488;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hyaluronan synthase 1 (Fragment).
GN Name=has1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed14729574;
 RA Bakkers J., Kramer C., Pothof J., Quaedvlieg N.E., Spink H.P.,
 RA Hamerschmidt M.;
 RT "Hae2 is required upstream of Rac1 to govern dorsal migration of
 RT lateral cells during zebrafish gastrulation.";
 RL Development 131:525-537(2004).
 DR EMBL; AY437407; AAR97372.1; -.
 DR ZFIN; ZDB-GENE-040218-2; hae1.
 DR InterPro; IPR001173; Glyco_trans_2.
 FT NON_TER 458 458
 SQ SEQUENCE 458 AA; 53188 MW; 4B421299ABBF96E CRC64;

Query Match 61.8%; Score 1048.5; DB 2; Length 458;
 Best Local Similarity 62.3%; Pred. No. 1.3e-81;
 Matches 195; Conservative 43; Mismatches 56; Indels 19; Gaps 4;

Qy 3 LSAHLVQSLPAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLROCLASARALL 62
 Db 57 LSLHVLIOQSFPAFVHGHNNARRK---PCSYTKTIGFTISAYQEDPAYLRECLQSVRALQ 113
 Qy 63 YPRALRLVMVDGNRAEDLYMVDMFREVFADDEPATYVWGNHYQWPEPAA---AGAVG 119
 Db 114 YPSELLRIVMVDGNSEDDRYMLENFRFVADQDGCYIWNQNYHS-WNPNQDQGAEG 172
 Qy 120 AGAYREVEAEDPGLRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGSDVDYVQVCDSD 179
 Db 173 PDADYEVLFEDPQRLEVEEIIIRTKVCIMQKWKREVMYTAFAKALGSSADFIQVCDSD 232
 Qy 180 TELDPMLLELVRVLDDEPRVGAAGVDVRIINPLDSWTSFLSSIRYVAFNVERACQSYF 239
 Db 233 TKLDPATLATELVCKLVESNOKYGAAGVDVMIINLKDYSISFMSLSLYWAFNVERACQSYF 292
 Qy 240 HCVSICISGLPLGYRNLLQQLFLEAWYQKFLGTHCTFTGDDRLHNTNRLSMGYATK---- 294
 Db 293 DCVCSICISGLPLGYRNLLQQLFLEAWYQKFLGTHCTFTGDDRLHNTNRLSMGYATK 352
 Qy 295 -----AEGTRW 300
 Db 353 KCYTETPAQFLRW 365

RESULT 8
 HASI_XENLA
 ID HASI_XENLA
 AC P13563;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
 DE Hyaluronan acid synthase 1 (HA synthase 1) (XHAS1) (DG42 protein).
 GN Name=HAS1;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8335;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88313363; PubMed=3410156;
 RA Rosa F., Sargent T.D., Rebert M.L., Michaels G.S., Jamrich M.,
 RA Grunz H., Jonas E., Winkles J.A., Dawid I.B.;
 RT "Accumulation and decay of DG42 gene products follow a gradient
 RT pattern during Xenopus embryogenesis.";
 RL Dev. Biol. 129:114-123 (1988).
 CC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
 CC synthesis (By similarity). May play a role in signaling or pattern
 CC formation in embryonic development.

CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
 CC glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-
 CC glucuronosyl(1->3)](n) + 2n UDP.
 CC -!- COFACTOR: Magnesium.
 CC -!- PATHWAY: Hyaluronate synthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- TISSUE SPECIFICITY: Embryo.
 CC -!- DEVELOPMENTAL STAGE: During gastrula and early neurula stages.
 CC -!- SIMILARITY: Belongs to the nodC/HAS family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M22249; AAA49699.1; -.
 DR PIR; A43740; A43740.
 DR InterPro; IPR001173; Glyco_trans_2.
 KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
 FT DOMAIN 1 28 Cytoplasmic (Potential).
 FT TRANSMEM 29 49 1 (Potential).
 FT DOMAIN 50 61 Extracellular (Potential).
 FT TRANSMEM 62 82 2 (Potential).
 FT DOMAIN 83 411 Cytoplasmic (Potential).
 FT TRANSMEM 412 432 3 (Potential).
 FT DOMAIN 433 433 Extracellular (Potential).
 FT TRANSMEM 434 454 4 (Potential).
 FT DOMAIN 455 456 Cytoplasmic (Potential).
 FT TRANSMEM 457 477 5 (Potential).
 FT DOMAIN 478 505 Extracellular (Potential).
 FT TRANSMEM 506 526 6 (Potential).
 FT DOMAIN 527 543 Cytoplasmic (Potential).
 FT TRANSMEM 544 564 7 (Potential).
 FT DOMAIN 565 588 Extracellular (Potential).
 SQ SEQUENCE 588 AA; 68522 MW; 33DA3B8E331F4CE9 CRC64;

Query Match 56.9%; Score 965.5; DB 1; Length 588;
 Best Local Similarity 60.5%; Pred. No. 2.4e-74;
 Matches 184; Conservative 47; Mismatches 62; Indels 11; Gaps 4;

Qy 6 HLVAQSLPAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLROCLASARALLYPR 65
 Db 72 HLMWQSLFAFLERVR---NKSLEPCSKTKVTAIAGYQENPEYLKCLBCKVYKPK 128
 Qy 66 ARLRVLMVDGNRAEDLYMVDMFREVFADDEPATYVWGNHYQWPEPAAAGVAGAYRE 125
 Db 129 DKLKIILVIDGNTDDAYMMEMFKDVFHGEDVGTVMKGNHYTVKKPE---ETNKGSCPE 185
 Qy 126 VEA---EDPGLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDTRL 182
 Db 186 VSKPLNDEGINMVEELVRNKRVCIMQWGGKREVMYTAFAIGTSVDYVQVCDSDTKL 245
 Qy 183 DPMALLELVRVLDDEPRVGAAGVDVRIINPLDSWTSFLSSIRYVAFNVERACQSYFHCV 242
 Db 246 DELATVENVKLESNDMYGAAGVDVRIINPLDYSFISFMSSLRVWAFNVERACQSYFDCV 305
 Qy 243 SCISGIPGLGYRNLLQQLFLEAWYQKFLGTHCTFTGDDRLHNTNRLSMGYATK--AEGTRW 300
 Db 306 SCISGIPGLGYRNLLQQLFLEAWYQKFLGTHCTFTGDDRLHNTNRLSMGYATK 365
 Qy 301 SGTP 304
 Db 366 SETP 369

RESULT 9
 Q95M29
 ID Q95M29 PRELIMINARY; PRT; 552 AA.
 AC Q95M29;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)


```
CC -!- SIMILARITY: Belongs to the nodC/HAS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; U52524; AAC53309.2; -.
CC DR EMBL; U69695; AAB17609.1; -.
CC DR EMBL; U53222; AAC52651.1; -.
CC DR MGI; 107821; Ha82.
CC DR InterPro; IPR001173; Glyco trans 2.
CC DR Pfam; PF00535; Glycos transf 2; 1.
CC KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
CC FT DOMAIN 1 11
CC FT TRANSMEM 12 32
CC FT DOMAIN 33 45
CC FT TRANSMEM 46 66
CC FT DOMAIN 67 374
CC FT TRANSMEM 375 395
CC FT DOMAIN 396 402
CC FT TRANSMEM 403 423
CC FT DOMAIN 424 429
CC FT TRANSMEM 430 450
CC FT DOMAIN 451 475
CC FT TRANSMEM 476 496
CC FT DOMAIN 497 510
CC FT TRANSMEM 511 531
CC FT DOMAIN 532 552
CC FT CONFLICT 138 138 I -> M (in Ref. 3).
CC SQ SEQUENCE 552 AA; 63492 MW; 7C88019F680DD982 CRC64;

Query Match 55.1%; Score 935; DB 1; Length 552;
Best Local Similarity 55.0%; Pred. No. 9.4e-72;
Matches 176; Conservative 50; Mismatches 60; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYOEDPAYLRQCILASARA 60
Db 51 AFLASHLIIQSLFAPLEHRRKMKKSLETPI--KLNKTVALCIAAYQEDPDYLRKCLQSVKR 108

Qy 61 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQPWPAPAAAGVGA 120
Db 109 LTPY--GIKVVWIDGNSDDLYMMDIFSEVIGRDKSATYIWKNNFHEK-----GP 157

Qy 121 GAYREVEAEDPGRLAVEALVTRRCVCAQWGGKREVMTAFKALGDSVDYVQVCDSDT 180
Db 158 GETESHKESQ--VTQLVLSNKSICIMQWGGKREVMTAFKALGDSVDYVQVCDSDT 215

Qy 181 RLDPMALLELRVLDDEPRVAGGVDVRLNPLDSWSFLSSRYVAFNVERACQSYFH 240
Db 216 MLDPASSVEMVKVLEEDPMVGCGDVQILNKYDSWISFLSSRYVAFNVERACQSYFG 275

Qy 241 CVCSISGPLGLYRNLLQOFLQLEAWYNOKFLGTHCTFGDDRHLTNRLMSGYATK----- 294
Db 276 CVQCISGPLGLMYRNSLLHEFVEDWYNQEFMGSCQSGFDDRLTNRLVSLGYATKYTARSK 335

Qy 295 -----ABGTRWS 301
Db 336 CLTETPIEYLRWLNLNQOTRWS 355

RESULT 15
ID Q95M16 PRELIMINARY; PRT; 552 AA.
AC Q95M16;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hyaluronic acid synthase 2.
GN Name=HAS2;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae;
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21363038; PubMed=11470161; DOI=10.1016/S0167-4781(01)00256-1;
RA Ohno S., Tanimoto K., Fujimoto K., Ijuin C., Honda K., Tanaka N.,
RA Doi T., Nakahara M., Tanne K.;
RT "Molecular cloning of rabbit hyaluronic acid synthases and their
RL Biochim. Biophys. Acta 1520:71-78 (2001).
DR EMBL; AB055978; BAB63264.1; -.
DR InterPro; IPR001173; Glyco trans 2.
DR SQ SEQUENCE 552 AA; 63483 MW; 4A1A656960B0B7B1 CRC64;

Query Match 55.1%; Score 935; DB 2; Length 552;
Best Local Similarity 55.3%; Pred. No. 9.4e-72;
Matches 177; Conservative 49; Mismatches 60; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYOEDPAYLRQCILASARA 60
Db 51 AFLASHLIIQSLFAPLEHRRKMKKSLETPI--KLNKTVALCIAAYQEDPDYLRKCLQSVKR 108

Qy 61 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQPWPAPAAAGVGA 120
Db 109 LTPY--GIKVVWIDGNSDDLYMMDIFSEVIGRDKSATYIWKNNFHEK-----GP 157

Qy 121 GAYREVEAEDPGRLAVEALVTRRCVCAQWGGKREVMTAFKALGDSVDYVQVCDSDT 180
Db 158 GETESHKESQ--VTQLVLSNKSICIMQWGGKREVMTAFKALGDSVDYVQVCDSDT 215

Qy 181 RLDPMALLELRVLDDEPRVAGGVDVRLNPLDSWSFLSSRYVAFNVERACQSYFH 240
Db 216 MLDPASSVEMVKVLEEDPMVGCGDVQILNKYDSWISFLSSRYVAFNVERACQSYFG 275

Qy 241 CVCSISGPLGLYRNLLQOFLQLEAWYNOKFLGTHCTFGDDRHLTNRLMSGYATK----- 294
Db 276 CVQCISGPLGLMYRNSLLHEFVEDWYNQEFMGSCQSGFDDRLTNRLVSLGYATKYTARSK 335

Qy 295 -----ABGTRWS 301
Db 336 CLTETPIEYLRWLNLNQOTRWS 355

RESULT 14
ID Q8HZJ3 PRELIMINARY; PRT; 552 AA.
AC Q8HZJ3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hyaluronan synthase 2.
OS Equus caballus (Horse).
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:11:23 ; Search time 23.3333 Seconds
(without alignments)
1023.759 Million cell updates/sec

Title: US-10-672-399-8
Perfect score: 1697
Sequence: 1 AFLSAHLVAOSLFAYLEHRR.....SGTTPGPAATQRRPRSCGG 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1553.5	91.5	578	US-08-635-552A-2	Sequence 2, Appli
2	1553.5	91.5	589	US-09-949-016-11530	Sequence 11530, A
3	1550.5	91.4	543	US-09-155-768-4	Sequence 4, Appli
4	1508.5	88.9	583	US-08-675-499A-3	Sequence 3, Appli
5	1508.5	88.9	583	US-08-812-008-3	Sequence 3, Appli
6	965.5	56.9	587	US-08-635-552A-3	Sequence 3, Appli
7	949	55.9	587	US-08-675-499A-4	Sequence 4, Appli
8	949	55.9	587	US-08-812-008-4	Sequence 4, Appli
9	935.5	55.1	554	US-08-812-008-32	Sequence 32, Appli
10	935	55.1	552	US-08-675-499A-2	Sequence 2, Appli
11	935	55.1	552	US-08-812-008-2	Sequence 2, Appli
12	934	55.0	552	US-09-949-016-6608	Sequence 6608, Ap
13	928	54.7	553	US-09-949-016-9599	Sequence 9599, Ap
14	415.5	24.5	241	US-08-865-273-2	Sequence 2, Appli
15	415.5	24.5	241	US-09-385-174-2	Sequence 2, Appli
16	409	24.1	568	US-09-469-200E-10	Sequence 10, Appli
17	279	16.4	426	US-08-675-499A-6	Sequence 6, Appli
18	279	16.4	426	US-08-812-008-6	Sequence 6, Appli
19	268.5	15.8	190	US-08-812-008-29	Sequence 29, Appli
20	268.5	15.8	190	US-08-812-008-30	Sequence 30, Appli
21	265	15.6	78	US-08-675-499A-24	Sequence 24, Appli
22	265	15.6	78	US-08-675-499A-27	Sequence 27, Appli
23	265	15.6	78	US-08-675-499A-28	Sequence 28, Appli
24	265	15.6	78	US-08-812-008-24	Sequence 24, Appli
25	265	15.6	78	US-08-812-008-27	Sequence 27, Appli
26	265	15.6	78	US-08-812-008-28	Sequence 28, Appli
27	231.5	13.6	417	US-09-469-200E-2	Sequence 2, Appli

28	217	12.8	43	4	US-08-675-499A-9	Sequence 9, Appli
29	217	12.8	43	4	US-08-812-008-9	Sequence 9, Appli
30	213.5	12.6	395	4	US-08-635-552A-4	Sequence 4, Appli
31	213.5	12.6	419	2	US-08-270-581-2	Sequence 2, Appli
32	213.5	12.6	419	4	US-09-146-893-2	Sequence 2, Appli
33	213.5	12.6	419	4	US-08-675-499A-5	Sequence 5, Appli
34	213.5	12.6	419	4	US-08-812-008-5	Sequence 5, Appli
35	177	10.4	43	4	US-08-812-008-35	Sequence 35, Appli
36	173	10.2	43	4	US-08-675-499A-7	Sequence 7, Appli
37	173	10.2	43	4	US-08-812-008-7	Sequence 7, Appli
38	163	9.6	43	4	US-08-675-499A-10	Sequence 10, Appli
39	163	9.6	43	4	US-08-812-008-10	Sequence 10, Appli
40	160	9.4	403	4	US-09-902-540-11529	Sequence 11529, A
41	153.5	9.0	1093	4	US-09-248-796A-17108	Sequence 17108, A
42	143.5	8.5	416	2	US-08-867-030B-12	Sequence 12, Appli
43	143.5	8.5	416	5	PCT-US95-06119-12	Sequence 12, Appli
44	133.5	7.9	393	4	US-09-902-540-11514	Sequence 11514, A
45	130.5	7.7	55	4	US-08-675-499A-29	Sequence 29, Appli

ALIGNMENTS

RESULT 1
US-08-635-552A-2
; Sequence 2, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS98-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-635-552A-2

Query Match	91.5%;	Score 1553.5;	DB 4;	Length 578;
Best Local Similarity	93.1%;	Pred. No. 4.9e-173;		
Matches 298;	Conservative 2;	Mismatches 1;	Indels 19;	Gaps 1;
Qy	1	AFLSAHLVAOSLFAYLEHRRVAAAARGPLODAATARSVALTISAYQEPAYLRQCCLASARA	60	
Db	61	AFLSAHLVAOSLFAYLEHRRVAAAARGPLODAATARSVALTISAYQEPAYLRQCCLASARA	120	
Qy	61	LLYPBRLRLVLMVDGNRAEDLYNVDMPREVFADDEPATYVWDQNYHQWPFAAGAVGA	120	

Db 121 LLYPRARLRVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQWEPAAAAAGVGA 180
Qy 121 GAYREVEADDPGRLAVALVTRRCVCAVQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 181 GAYREVEADDPGRLAVALVTRRCVCAVQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Qy 181 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 240
Db 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 300
Qy 241 CVSCISGPLGLYRNLLQQFLAEAWYKFLGTHCTFGDDRHLTNRMLSMGVATK----- 294
Db 301 CVSCISGPLGLYRNLLQQFLAEAWYKFLGTHCTFGDDRHLTNRMLSMGVATKTSRSR 360
Qy 295 -----AEGTRWS 301
Db 361 CYSETPSSFLRWLSQQTRWS 380

RESULT 2

US-09-949-016-11530
; Sequence 11530, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASES, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11530
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11530

Query Match 91.5%; Score 1553.5; DB 4; Length 589;
Best Local Similarity 93.1%; Pred. No. 5.1e-173;
Matches 298; Conservative 2; Mismatches 1; Indels 19; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 72 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 131
Qy 61 LLYPRARLRVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQWEPAAAAAGVGA 120
Db 132 LLYPRARLRVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQWEPAAAAAGVGA 191
Qy 121 GAYREVEADDPGRLAVALVTRRCVCAVQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 192 GAYREVEADDPGRLAVALVTRRCVCAVQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 251
Qy 181 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 240
Db 252 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 311
Qy 241 CVSCISGPLGLYRNLLQQFLAEAWYKFLGTHCTFGDDRHLTNRMLSMGVATK----- 294
Db 312 CVSCISGPLGLYRNLLQQFLAEAWYKFLGTHCTFGDDRHLTNRMLSMGVATKTSRSR 371
Qy 295 -----AEGTRWS 301
Db 372 CYSETPSSFLRWLSQQTRWS 391

RESULT 3

US-09-155-768-4
; Sequence 4, Application US/09155768A
; Patent No. 6162908
; GENERAL INFORMATION:
; APPLICANT: SEIKAGAKU CORPORATION
; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF
HUMAN-ORIGIN AND DNA ENCODING THE SAME
; FILE REFERENCE: TOYAM37,001APC
; CURRENT APPLICATION NUMBER: US/09/155,768A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: JP 8-084326
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: JP 8-109663
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 543
; TYPE: PRT
; ORGANISM: HUMAN
US-09-155-768-4

Query Match 91.4%; Score 1550.5; DB 3; Length 543;
Best Local Similarity 92.8%; Pred. No. 1e-172;
Matches 297; Conservative 3; Mismatches 1; Indels 19; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 26 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 85
Qy 61 LLYPRARLRVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQWEPAAAAAGVGA 120
Db 86 LLYPRARLRVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQWEPAAAAAGVGA 145
Qy 121 GAYREVEADDPGRLAVALVTRRCVCAVQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 146 GAYREVEADDPGRLAVALVTRRCVCAVQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 205
Qy 181 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 240
Db 206 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 265
Qy 241 CVSCISGPLGLYRNLLQQFLAEAWYKFLGTHCTFGDDRHLTNRMLSMGVATK----- 294
Db 266 CVSCISGPLGLYRNLLQQFLAEAWYKFLGTHCTFGDDRHLTNRMLSMGVATKTSRSR 325
Qy 295 -----AEGTRWS 301
Db 326 CYSETPSSFLRWLSQQTRWS 345

RESULT 4

US-08-675-499A-3
; Sequence 3, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/675,499A
FILING DATE: 03-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 150.170U51
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-359-3260
TELEFAX: 612-359-3263
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-675-499A-3

Query Match 88.9%; Score 1508.5; DB 4; Length 583;
Best Local Similarity 89.9%; Pred. No. 9.5e-168;
Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYVMDMREVFADDPATYVWDGNHYQHPWEPA-A 114
DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYVMDMREVFADDPATYVWDGNHYQHPWEPA 179
QY 115 AGAVGAGAYREVEADDPGLRAVEALVTRRCVCVQAQRWGGKREVMTAFKALGDSVDYVQ 174
DB 180 TGAVEGAYREVEADDPGLRAVEALVTRRCVCVQAQRWGGKREVMTAFKALGDSVDYVQ 239
QY 175 VCDSDTRLDPMALLELRVLDDEPRVAGVGGDVRILNPLDSWVSLSLRYWVAFNVERA 234
DB 240 VCDSDTRLDPMALLELRVLDDEPRVAGVGGDVRILNPLDSWVSLSLRYWVAFNVERA 299
QY 235 COSYFHCVCISGPIGLYRNLLQOFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 294
DB 300 COSYFHCVCISGPIGLYRNLLQOFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 359
QY 295 -----AEGTRWS 301
DB 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 5
US-08-812-008-3
Sequence 3, Application US/08812008
Patent No. 6602693
GENERAL INFORMATION:
APPLICANT: McDonald, J. A.
APPLICANT: Spicer, A. P.
APPLICANT: Augustine, M. L.
TITLE OF INVENTION: GENE ENCODING HYALURONAN
TITLE OF INVENTION: SYNTHASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: U.S.A
ZIP: 55402

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,008
FILING DATE: 05-MAR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/575,499
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Embretson, Janet E
REGISTRATION NUMBER: 39,665
REFERENCE/DOCKET NUMBER: 150.183U51
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-812-008-3

Query Match 88.9%; Score 1508.5; DB 4; Length 583;
Best Local Similarity 89.9%; Pred. No. 9.5e-168;
Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYVMDMREVFADDPATYVWDGNHYQHPWEPA-A 114
DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYVMDMREVFADDPATYVWDGNHYQHPWEPA 179
QY 115 AGAVGAGAYREVEADDPGLRAVEALVTRRCVCVQAQRWGGKREVMTAFKALGDSVDYVQ 174
DB 180 TGAVEGAYREVEADDPGLRAVEALVTRRCVCVQAQRWGGKREVMTAFKALGDSVDYVQ 239
QY 175 VCDSDTRLDPMALLELRVLDDEPRVAGVGGDVRILNPLDSWVSLSLRYWVAFNVERA 234
DB 240 VCDSDTRLDPMALLELRVLDDEPRVAGVGGDVRILNPLDSWVSLSLRYWVAFNVERA 299
QY 235 COSYFHCVCISGPIGLYRNLLQOFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 294
DB 300 COSYFHCVCISGPIGLYRNLLQOFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 359
QY 295 -----AEGTRWS 301
DB 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 6
US-08-635-552A-3
Sequence 3, Application US/08635552A
Patent No. 6423514
GENERAL INFORMATION:
APPLICANT: Briekin, Michael J.
TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
TITLE OF INVENTION: Acids and Uses Thereof
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA

```
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-635-552A-3
```

```
Query Match 56.9%; Score 965.5; DB 4; Length 587;
Best Local Similarity 60.5%; Pred. No. 5.3e-104;
Matches 184; Conservative 47; Mismatches 62; Indels 11; Gaps 4;

Qy 6 HLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 HLMQSLFAFLAIRRV---NKSELPKFKKVTALTIAGYQENPEYLIKLESCKVVKPK 128

Qy 66 ARLRVLMVDGNRAEDLYWDMFREVFADEDPATYVWDGNHQWPEAAAGAGAGAYRE 125
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 DKLKILVIDGNTEDDAYMMEMFKDVFHGEDVGTVMKGNVHTVKKPE---ETNKGSCPE 185

Qy 126 VEA---EDPGLAVEALVTRRCVCAQRMGKREVMTAFKALGDSVDYVQVCDSTRL 182
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 VSKPLNEDEGINWVEELVRNKRVCVCIQWQWGGKREVMTAFQAIGTSVDYVQVCDSTKL 245

Qy 183 DPMALLELRVLDEDPRVGAGVDVRIINPLDSDWVSFLSLRYWVAFNVERACOSYFHCY 242
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
246 DELATVEMVKVLESNDMTYGAVGVDVRIINLPYDSFISFMSLRVWMAFNVERACOSYFDCV 305

Qy 243 SCISGPLGLYRNLLIQFLEAWYQKFLGTCTFGDDRHLLTNRLSMGYATK--AEGTRW 300
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 SCISGPLGMRYNNILQVLEAWYRQKFLGTCTGLGDDRHLLTNRLSMGYRTKYTHKSRAP 365

Qy 301 SGTP 304
Db |||
366 SETP 369
```

RESULT 7

```
US-08-675-499A-4
; Sequence 4, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwesman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-499A-4
```

```
Query Match 55.9%; Score 949; DB 4; Length 587;
Best Local Similarity 60.2%; Pred. No. 4.6e-102;
Matches 183; Conservative 47; Mismatches 62; Indels 12; Gaps 5;

Qy 6 HLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 HLMQSLFAFLAIRRV---NKSELPKFKKVTALTIAGYQENPEYLIKLESCKVVKPK 128

Qy 66 ARLRVLMVDGNRAEDLYWDMFREVFADEDPATYVWDGNHQWPEAAAGAGAGAYRE 125
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 DKLKILVIDGNTEDDAYMMEMFKDVFHGEDVGTVMKGNVHTVKKPE---ETNKGSCPE 185

Qy 126 VEA---EDPGLAVEALVTRRCVCAQRMGKREVMTAFKALGDSVDYVQVCDSTRL 182
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 VSKPLNEDEGINWVEELVRNKRVCVCIQWQWGGKREVMTAFQAIGTSVDYVQVCDSTKL 244

Qy 183 DPMALLELRVLDEDPRVGAGVDVRIINPLDSDWVSFLSLRYWVAFNVERACOSYFHCY 242
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
245 DELATVEMVKVLESNDMTYGAVGVDVRIINLPYDSFISFMSLRVWMAFNVERACOSYFDCV 304

Qy 243 SCISGPLGLYRNLLIQFLEAWYQKFLGTCTFGDDRHLLTNRLSMGYATK--AEGTRW 300
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
305 SCISGPLGMRYNNILQVLEAWYRQKFLGTCTGLGDDRHLLTNRLSMGYRTKYTHKSRAP 364

Qy 301 SGTP 304
Db |||
365 SETP 368
```

RESULT 8

```
US-08-612-008-4
; Sequence 4, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwesman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
```



```
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-675-499A-2

Query Match 55.1%; Score 935; DB 4; Length 552;
Best Local Similarity 55.0%; Pred. No. 1.8e-100;
Matches 176; Conservative 50; Mismatches 60; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 51 AFLASHLIIQSLFAPLEHRRKMKKSLETPI--KLKNTVALCIAAYQEDPDYLRKCLQSVKR 108
Qy 61 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPWEPAAAGAVGA 120
Db 109 LTYP--GIKVMVVDGNSDDLLYMDIFSEVIGRDKSATYIWKKNFHEK-----GP 157
Qy 121 GAYREVEAEDPGRLAVALVTRRCVCAQWKGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 158 GETEESHKESQH--VTQLVLSNKSICIMQKMGKREVMYTAFAKALGDSVDYVQVCDSDT 215
Qy 181 RLDPMALLELRVLDEDPYRGVAGDVRILNPLDSWVSFLSSLYRYWAFNVERACQSYFH 240
Db 216 MLDPASSVEMVKVLEEDPMVGGVDVQILNKYDSWISFLSSVRYWMAFNIERACQSYFG 275
Qy 241 CVCSIGPLGLYRNLLQOFLWYKQFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
Db 276 CVQCISGPLGMYRNSLLHEFVEDWYNOBFMGNCQSGFDDRLTNRLVSLGYATKYTARSK 335
Qy 295 -----AEGTRWS 301
Db 336 CLTETPIEYLRWLNQOTRWS 355

RESULT 11
US-08-812-008-2
; Sequence 2, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:

; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustine, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; CORRESPONDENCE ADDRESS:
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-812-008-2

Query Match 55.1%; Score 935; DB 4; Length 552;
Best Local Similarity 55.0%; Pred. No. 1.8e-100;
Matches 176; Conservative 50; Mismatches 60; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 51 AFLASHLIIQSLFAPLEHRRKMKKSLETPI--KLKNTVALCIAAYQEDPDYLRKCLQSVKR 108
Qy 61 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPWEPAAAGAVGA 120
Db 109 LTYP--GIKVMVVDGNSDDLLYMDIFSEVIGRDKSATYIWKKNFHEK-----GP 157
Qy 121 GAYREVEAEDPGRLAVALVTRRCVCAQWKGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 158 GETEESHKESQH--VTQLVLSNKSICIMQKMGKREVMYTAFAKALGDSVDYVQVCDSDT 215
Qy 181 RLDPMALLELRVLDEDPYRGVAGDVRILNPLDSWVSFLSSLYRYWAFNVERACQSYFH 240
Db 216 MLDPASSVEMVKVLEEDPMVGGVDVQILNKYDSWISFLSSVRYWMAFNIERACQSYFG 275
Qy 241 CVCSIGPLGLYRNLLQOFLWYKQFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
Db 276 CVQCISGPLGMYRNSLLHEFVEDWYNOBFMGNCQSGFDDRLTNRLVSLGYATKYTARSK 335
Qy 295 -----AEGTRWS 301
Db 336 CLTETPIEYLRWLNQOTRWS 355

RESULT 12
US-09-949-016-6608
```

```
; Sequence 6608, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6608
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6608

Query Match      55.0%; Score 934; DB 4; Length 552;
Best Local Similarity 55.0%; Pred. No. 2.4e-100;
Matches 176; Conservative 50; Mismatches 60; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 51 AFLASHLLIQSLFAPLEHRRKWKSLTFI--KLNTKVALCIAAQVEDPDYLRKCLQSVKR 108

Qy 61 LLYPRARLRVLMVDGNRAEDLYMDMFRFVFADEDPATYVMDGNHQPWEPAAAGAVGA 120
Db 109 LTYP--GIKVVWVIDGNEDDLDMYMDIFSEVWGRDKSATYIWKNNFHEK-----GP 157

Qy 121 GAYREVEADPGRLAVALRTRRCVCVQAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 158 GETDESHKESQSH--VTQLVLSNKSICIMQKGGKREVMYTAFAKALGSRVDYVQVCDSDT 215

Qy 181 RLDPMLLELVRVLDEDPVGAAGVDVRLINPLDSWVSFLSLRYWVAFNVERACOSYFH 240
Db 216 MLDPASSVEMWVLEEDPMVGGVDVQILNKYDSWISFLSSVRYWMAFNVERACOSYFG 275

Qy 241 CVSCISGPIGLYRNLLQOFLBANYNQKPLGTHCTFGDDRHLTNRLMSMGYATK----- 294
Db 276 CVQCISGPIGLMYRNSLLHEFVEDWYNOEFMGNCQCSFGDDRHLTNRLVSLGYATKYTARSK 335

Qy 295 -----ABGTRWS 301
Db 336 CLTETPIEYLRWLNQOTRWS 355

RESULT 13
US-09-949-016-9599
; Sequence 9599, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9599
; LENGTH: 553
```

```
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9599

Query Match      54.7%; Score 928; DB 4; Length 553;
Best Local Similarity 56.8%; Pred. No. 1.2e-99;
Matches 183; Conservative 40; Mismatches 65; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASAR 59
Db 50 AILGLHLIIQSLFAPLEHRRMRAGQALKJPSPRGVSVALCIAAYQEDPDYLRKCLRAQ 109

Qy 60 ALLYPRARLRVLMVDGNRAEDLYMDMFRFVFADEDPATYVMDGNHQPWEPAAAGAV 118
Db 110 RISFP--DLKVVWVVDGNRQEDAYMLDIFHEVLGGTEQAQGFVWRSNFHEAGEGETEASL 167

Qy 119 GAGAYREVEADPGRLAVALRTRRCVCVQAQWGGKREVMYTAFAKALGDSVDYVQVCDSD 178
Db 168 QEGMDR-----VRDVVRASTFCINQKGGKREVMYTAFAKALGDSVDYVQVCDSD 216

Qy 179 DTRLDPMLLELVRVLDEDPVGAAGVDVRLINPLDSWVSFLSLRYWVAFNVERACOSY 238
Db 217 DTVLDPACTIEMLRVLEEDPMVGGVDVQILNKYDSWISFLSSVRYWMAFNVERACOSY 276

Qy 239 FHCVCISGPIGLYRNLLQOFLBANYNQKPLGTHCTFGDDRHLTNRLMSMGYATK---- 294
Db 277 FGCVCISGPIGLMYRNSLLQOFLBANYNQKPLGTHCTFGDDRHLTNRLVSLGYATKYTAR 336

Qy 295 -----ABGTRWS 301
Db 337 SKCLTETPTKYLRWLNQOTRWS 358

RESULT 14
US-08-865-273-2
; Sequence 2, Application US/08865273
; Patent No. 5994100
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN
; APPLICANT: NAMBI, PONNAL
; APPLICANT: PULLEN, MARK A
; TITLE OF INVENTION: NOVEL HAS2 SPLICING VARIANT
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,273
; FILING DATE: 29-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```

INFLAMMAT

LENGTH: 241 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-865-273-2

Query Match 24.5%; Score 415.5; DB 2; Length 241;
Best Local Similarity 48.0%; Pred. No. 4.7e-40;
Matches 84; Conservative 31; Mismatches 45; Indels 15; Gaps 4;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 51 AFLASHLIIQSLFAPLEHRRKMKKSLETPI--KLNTVALCIAAYQEDPDYLRKCLQSVKR 108
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 61 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQHPWEPAAAGAVGA 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 109 LTYP--GIKVVWIDGNSEDDLYMMDIFSEVMGRDKSATHIWKNFHEK-----GP 157
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 121 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGKREVMYTAFAKALGDSVDYVQV 175
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 158 GETDESHKESQH--VTQLVLSNKSICIMQKMGKREVMYTAFRALGRSVDYVQV 210
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Search completed: March 11, 2005, 14:24:38
Job time : 24.3333 secs

LENGTH: 241 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-865-273-2

Query Match 24.5%; Score 415.5; DB 2; Length 241;
Best Local Similarity 48.0%; Pred. No. 4.7e-40;
Matches 84; Conservative 31; Mismatches 45; Indels 15; Gaps 4;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 51 AFLASHLIIQSLFAPLEHRRKMKKSLETPI--KLNTVALCIAAYQEDPDYLRKCLQSVKR 108
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 61 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQHPWEPAAAGAVGA 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 109 LTYP--GIKVVWIDGNSEDDLYMMDIFSEVMGRDKSATHIWKNFHEK-----GP 157
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 121 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGKREVMYTAFAKALGDSVDYVQV 175
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 158 GETDESHKESQH--VTQLVLSNKSICIMQKMGKREVMYTAFRALGRSVDYVQV 210
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 15

US-09-385-174-2

; Sequence 2, Application US/09385174

; Patent No. 6350446

; GENERAL INFORMATION:

; APPLICANT: ZHU, YUAN

; NAMBI, PONNAL

; PULLEN, MARK A

; TITLE OF INVENTION: NOVEL HAS2 SPLICING VARIANT

; H0EFC11: A TARGET IN CHRONIC RENAL FAILURE, INFLAMMATORY DISEASES AND MYOCARDIAL ISCHEMIA

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P.O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/385,174

; FILING DATE: 30-Aug-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/865,273

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: PRESTIA, PAUL F

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: GH-70053

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 241 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-385-174-2

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:22:15 ; Search time 71.8182 Seconds
(without alignments)
1469.688 Million cell updates/sec

Title: US-10-672-399-8
Perfect score: 1697
Sequence: 1 AFLSHLVQSLFAYLEHRR.....SGTPGPAATQRRPSPCGG 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1697	100.0	320	16	US-10-672-399-8
2	1553.5	91.5	578	13	US-10-042-523-2
3	1553.5	91.5	578	16	US-10-672-399-2
4	1508.5	88.9	583	10	US-09-902-939-4
5	1363	80.3	376	16	US-10-672-399-4
6	1298	78.5	360	16	US-10-672-399-6
7	965.5	56.9	587	13	US-10-042-523-3
8	965.5	56.9	588	15	US-10-309-560-10
9	935.5	55.1	554	10	US-09-902-939-6
10	935	55.1	552	10	US-09-902-939-5
11	935	55.1	552	14	US-10-262-526-4
12	934	55.0	552	14	US-10-262-526-2
13	928	54.7	553	15	US-10-295-027-370

14	928	54.7	553	15	US-10-188-832-137	Sequence 137, Appl
15	409	24.1	568	14	US-10-011-768B-10	Sequence 10, Appl
16	409	24.1	568	14	US-10-011-771B-10	Sequence 10, Appl
17	401.5	23.7	567	9	US-09-879-959-7	Sequence 7, Appl
18	401.5	23.7	567	14	US-10-172-527-7	Sequence 7, Appl
19	401.5	23.7	567	15	US-10-309-560-12	Sequence 12, Appl
20	233.5	13.8	417	15	US-10-309-560-19	Sequence 16, Appl
21	233.5	13.8	417	15	US-10-309-560-16	Sequence 19, Appl
22	233.5	13.8	417	15	US-10-309-560-22	Sequence 22, Appl
23	233.5	13.8	417	15	US-10-309-560-23	Sequence 23, Appl
24	233.5	13.8	417	15	US-10-309-560-25	Sequence 25, Appl
25	233.5	13.8	417	15	US-10-309-560-27	Sequence 27, Appl
26	233.5	13.8	417	15	US-10-309-560-28	Sequence 28, Appl
27	233.5	13.8	417	15	US-10-309-560-29	Sequence 29, Appl
28	231.5	13.6	417	9	US-09-879-959-2	Sequence 2, Appl
29	231.5	13.6	417	14	US-10-011-768B-2	Sequence 2, Appl
30	231.5	13.6	417	14	US-10-011-771B-2	Sequence 2, Appl
31	231.5	13.6	417	14	US-10-172-527-2	Sequence 2, Appl
32	231.5	13.6	417	14	US-10-326-185-2	Sequence 2, Appl
33	231.5	13.6	417	15	US-10-309-560-2	Sequence 2, Appl
34	231.5	13.6	417	15	US-10-309-560-15	Sequence 15, Appl
35	231.5	13.6	417	15	US-10-309-560-17	Sequence 17, Appl
36	231.5	13.6	417	15	US-10-309-560-18	Sequence 18, Appl
37	231.5	13.6	417	15	US-10-309-560-20	Sequence 20, Appl
38	231.5	13.6	417	15	US-10-309-560-21	Sequence 21, Appl
39	231.5	13.6	417	15	US-10-309-560-24	Sequence 24, Appl
40	231.5	13.6	417	15	US-10-309-560-26	Sequence 26, Appl
41	216.5	12.8	419	15	US-10-309-560-37	Sequence 37, Appl
42	216.5	12.8	419	15	US-10-309-560-51	Sequence 51, Appl
43	216.5	12.8	419	15	US-10-309-560-55	Sequence 55, Appl
44	216.5	12.8	419	15	US-10-309-560-56	Sequence 56, Appl
45	216.5	12.8	419	15	US-10-309-560-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-10-672-399-8
; Sequence 8, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-8

Query Match 100.0%; Score 1697; DB 16; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.6e-168;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA	60
DB	1	AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA	60
QY	61	LLYPRARLRLVMVDGNRAEDLYVMDMPREVFADDPATYVMDGNHGFWEPAAGVGA	120
DB	61	LLYPRARLRLVMVDGNRAEDLYVMDMPREVFADDPATYVMDGNHGFWEPAAGVGA	120
QY	121	GAYREVEAEDPGRLAVEALVETRRVCVCAQWGGKREVMYTAFAKLGSDVYVQVCDSDT	180
DB	121	GAYREVEAEDPGRLAVEALVETRRVCVCAQWGGKREVMYTAFAKLGSDVYVQVCDSDT	180

QY 181 RLDPMALLELRVLDEDPVGAAGVGDVRIINPLDSWVSFLSSLRVWVAFNVERACQSYFH 240
DB 181 RLDPMALLELRVLDEDPVGAAGVGDVRIINPLDSWVSFLSSLRVWVAFNVERACQSYFH 240
QY 241 CVSCISGPLGLYRNLLQQFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGYATKAEGRTRW 300
DB 241 CVSCISGPLGLYRNLLQQFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGYATKAEGRTRW 300
QY 301 SGTPPGAATQRRPRPSCGG 320
DB 301 SGTPPGAATQRRPRPSCGG 320

RESULT 2

US-10-042-523-2
; Sequence 2, Application US/10042523
; Publication No. US2002015026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LXS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-042-523-2
Query Match 91.5%; Score 1553.5; DB 13; Length 578;
Best Local Similarity 93.1%; Pred. No. 3.6e-153;
Matches 298; Conservative 2; Mismatches 1; Indels 19; Gaps 1;
QY 1 AFLSAHLVAQSLPAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCASARA 60
DB 61 AFLSAHLVAQSLPAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCASARA 120
QY 61 LLYPRARLRVLMVVDGNRAEDLYVMDVRFVFADEDPATYVWDGNYHQPWEPAAAGAVGA 120
DB 121 LLYPRARLRVLMVVDGNRAEDLYVMDVRFVFADEDPATYVWDGNYHQPWEPAAAGAVGA 180
QY 121 GAYREVEAEDPGRLAVALVTRRCVCVCAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 180
DB 181 GAYREVEAEDPGRLAVALVTRRCVCVCAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 240

QY 181 RLDPMALLELRVLDEDPVGAAGVGDVRIINPLDSWVSFLSSLRVWVAFNVERACQSYFH 240
DB 241 RLDPMALLELRVLDEDPVGAAGVGDVRIINPLDSWVSFLSSLRVWVAFNVERACQSYFH 300
QY 241 CVSCISGPLGLYRNLLQQFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
DB 301 CVSCISGPLGLYRNLLQQFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGYATKYTSRSR 360
QY 295 -----AEGTRWS 301
DB 361 CYSETPSSFLRWLSQQTRWS 380

RESULT 3

US-10-672-399-2
; Sequence 2, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-672-399-2

Query Match 91.5%; Score 1553.5; DB 16; Length 578;
Best Local Similarity 93.1%; Pred. No. 3.6e-153;
Matches 298; Conservative 2; Mismatches 1; Indels 19; Gaps 1;

QY 1 AFLSAHLVAQSLPAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCASARA 60
DB 61 AFLSAHLVAQSLPAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCASARA 120
QY 61 LLYPRARLRVLMVVDGNRAEDLYVMDVRFVFADEDPATYVWDGNYHQPWEPAAAGAVGA 120
DB 121 LLYPRARLRVLMVVDGNRAEDLYVMDVRFVFADEDPATYVWDGNYHQPWEPAAAGAVGA 180
QY 121 GAYREVEAEDPGRLAVALVTRRCVCVCAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 180
DB 181 GAYREVEAEDPGRLAVALVTRRCVCVCAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 240
QY 181 RLDPMALLELRVLDEDPVGAAGVGDVRIINPLDSWVSFLSSLRVWVAFNVERACQSYFH 240
DB 241 RLDPMALLELRVLDEDPVGAAGVGDVRIINPLDSWVSFLSSLRVWVAFNVERACQSYFH 300
QY 241 CVSCISGPLGLYRNLLQQFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
DB 301 CVSCISGPLGLYRNLLQQFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGYATKYTSRSR 360
QY 295 -----AEGTRWS 301
DB 361 CYSETPSSFLRWLSQQTRWS 380

RESULT 4

US-09-902-939-4
; Sequence 4, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHarya
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/0H020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: mouse
US-09-902-939-4

Query Match      88.9%; Score 1508.5; DB 10; Length 583;
Best Local Similarity 89.9%; Pred. No. 1.8e-148;
Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
Db 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
Qy 56 ASARALLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPA-A 114
Db 120 TSARALLYPHTRLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPA 179
Qy 115 AGAVGAGAYREVEADPGRLAVEALVTRRCVCVQAQRWGKGKREVMYTAFAKALGSDVDYVQ 174
Db 180 TGAVGEGAYREVEADPGRLAVEALVTRRCVCVQAQRWGKGKREVMYTAFAKALGSDVDYVQ 239
Qy 175 VCDSDTRLDPMALLELVRLVLDDEPRVGAAGDVRILNPLDSWVSLSLRYVWAFNVERA 234
Db 240 VCDSDTRLDPMALLELVRLVLDDEPRVGAAGDVRILNPLDSWVSLSLRYVWAFNVERA 299
Qy 235 CQSYFHCVCSISGPIGLYRNLLQOFLFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 294
Db 300 CQSYFHCVCSISGPIGLYRNLLQOFLFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 359
Qy 295 -----AEGTRWS 301
Db 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 5
US-10-672-399-4
; Sequence 4, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-4

Query Match      80.3%; Score 1363; DB 16; Length 376;
Best Local Similarity 83.4%; Pred. No. 1.6e-133;
Matches 267; Conservative 0; Mismatches 1; Indels 52; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Qy 61 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPAAGAVGA 120
Db 121 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPAAGAVGA 180
Qy 121 GAYREVEADPGRLAVEALVTRRCVCVQAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 180
Db 181 GAYREVEADPGRLAVEALVTRRCVCVQAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 240
Qy 181 GAYREVEADPGRLAVEALVTRRCVCVQAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 240
Db 240 GAYREVEADPGRLAVEALVTRRCVCVQAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 300

RESULT 6
US-10-672-399-6
; Sequence 6, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-6

Query Match      76.5%; Score 1298; DB 16; Length 360;
Best Local Similarity 95.1%; Pred. No. 9.7e-127;
Matches 252; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Qy 61 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPAAGAVGA 120
Db 121 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPAAGAVGA 180
Qy 121 GAYREVEADPGRLAVEALVTRRCVCVQAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 180
Db 181 GAYREVEADPGRLAVEALVTRRCVCVQAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 240
Qy 181 GAYREVEADPGRLAVEALVTRRCVCVQAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 240
Db 241 GAYREVEADPGRLAVEALVTRRCVCVQAQRWGKGKREVMYTAFAKALGSDVDYVQVCDSDT 300

RESULT 7
US-10-042-523-3
; Sequence 3, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; CORRESPONDENCE ADDRESS:
; NUMBER OF SEQUENCES: 4
; ADDRESS: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
```

```
;
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-042-523-3

Query Match 56.9%; Score 965.5; DB 13; Length 587;
Best Local Similarity 60.5%; Pred. No. 1.1e-91;
Matches 184; Conservative 47; Mismatches 62; Indels 11; Gaps 4;

Qy 6 HLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db 72 HLMQSLFAPLEIRRV---NKSELPSPKTKTVALTIAGYQENPEYLIKLCLESKYVYKPK 128

Qy 66 ARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNTHQHPWEPAAGAGAGAYRE 125
Db 129 DKLKILVIDGNTDDAYMMEMFKDVFHGEDVGYVMKGNHYHTVKKPE---ETNKGSCPE 185

Qy 126 VEA---EDPGLAVEALVTRRCVCVAORMGKREVMYTAFAKALGDSVDYVQVCDSDTRL 182
Db 186 VSKPLNEDEGINMVEELVRNRCVCINQMKGKREVMYTAFAQALGTSVDYVQVCDSDTKL 245

Qy 183 DPMALLELRVLDEDPRVGAGVDVRLINPLDLSWVSFLSSRLRYWVAFNVERACQSYFHCV 242
Db 246 DELATVEMVKVLESNDMYGAGVDVRLINPYDSFISFMSSRLRYWMAFNVERACQSYFDCV 305

Qy 243 SCISGPLGLYRNLLQOFLBANYNQKFLGTHCTFGDDRHLTNRLMSMGYATK--ABGTRW 300
Db 306 SCISGPLGMRYNNILQVFLBANYRQKFLGTCTGLGDDRHLNRLVLSMGYRTKYTHKGRAP 365

Qy 301 SGTP 304
Db 366 SETP 369

RESULT 8
US-10-309-560-10
; Sequence 10, Application US/10309560
; Publication No. US20030235893A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: KUMARI, KSHAMA
; TITLE OF INVENTION: HYALURONAN SYNTHASES AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 35541.082
; CURRENT APPLICATION NUMBER: US/10/309,560
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/336,105
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
```

```
;
;
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-309-560-10

Query Match 56.9%; Score 965.5; DB 15; Length 588;
Best Local Similarity 60.5%; Pred. No. 1.1e-91;
Matches 184; Conservative 47; Mismatches 62; Indels 11; Gaps 4;

Qy 6 HLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db 72 HLMQSLFAPLEIRRV---NKSELPSPKTKTVALTIAGYQENPEYLIKLCLESKYVYKPK 128

Qy 66 ARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNTHQHPWEPAAGAGAGAYRE 125
Db 129 DKLKILVIDGNTDDAYMMEMFKDVFHGEDVGYVMKGNHYHTVKKPE---ETNKGSCPE 185

Qy 126 VEA---EDPGLAVEALVTRRCVCVAORMGKREVMYTAFAKALGDSVDYVQVCDSDTRL 182
Db 186 VSKPLNEDEGINMVEELVRNRCVCINQMKGKREVMYTAFAQALGTSVDYVQVCDSDTKL 245

Qy 183 DPMALLELRVLDEDPRVGAGVDVRLINPLDLSWVSFLSSRLRYWVAFNVERACQSYFHCV 242
Db 246 DELATVEMVKVLESNDMYGAGVDVRLINPYDSFISFMSSRLRYWMAFNVERACQSYFDCV 305

Qy 243 SCISGPLGLYRNLLQOFLBANYNQKFLGTHCTFGDDRHLTNRLMSMGYATK--ABGTRW 300
Db 306 SCISGPLGMRYNNILQVFLBANYRQKFLGTCTGLGDDRHLNRLVLSMGYRTKYTHKGRAP 365

Qy 301 SGTP 304
Db 366 SETP 369

RESULT 9
US-09-902-939-6
; Sequence 6, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 554
; TYPE: PRT
; ORGANISM: mouse
US-09-902-939-6

Query Match 55.1%; Score 935.5; DB 10; Length 554;
Best Local Similarity 57.1%; Pred. No. 1.4e-88;
Matches 185; Conservative 41; Mismatches 61; Indels 37; Gaps 6;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLD---AATARSVALTISAYQEDPAYLRQCLAS 57
Db 50 AILGLHLLLIQSLFAPLEHREMRAGR-PLKLHCSQRSRSVALCIAAQEDPEYLRKCLRS 108

Qy 58 ARALLYPRARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNTHQHPWEPAAG 116
Db 109 AQRIAPF--NLKVMVMVDGNRQEDTYMLDIFHEVLGTEQAGFPFVMSNFHEAGEGTEA 166

Qy 117 AVGAGAYREVEADPCGLAVEALVTRRCVCVAORMGKREVMYTAFAKALGDSVDYVQVC 176
Db 167 SLOEGMER-----VRAVVMASTFSCIMQKMGKREVMYTAFAKALGNSVDYIQVC 215

Qy 177 DSDTRLDPMALLELRVLDEDPRVGAGVDVRLINPLDLSWVSFLSSRLRYWVAFNVERACQ 236
```


QY 295 -----AEGTRWS 301
Db 337 SKCLTETPTKYLRWLNQOTRWS 358

RESULT 15
US-10-011-768B-10
; Sequence 10, Application US/10011768B
; Publication No. US20030073221A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/10/011,768B
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Paramaecium bursaria chorella virus
US-10-011-768B-10

Query Match 24.1%; Score 409; DB 14; Length 568;
Best Local Similarity 32.9%; Pred. No. 1.7e-33;
Matches 103; Conservative 63; Mismatches 105; Indels 42; Gaps 10;

QY 2 FLSAHLVAQSLFAYLEHRR-----VAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 57
Db 56 FVFGFPLAQVLFSSELNRKLRKWLSPKGMNDV----RLAVIIAGYREDPTFMFOKCLAS 111

QY 58 ARALLYPRARLRVLMVDGNRAEDLYMDMFREVFEDPADPYVWDGNYHQWEPAPAAAGA 117
Db 112 VRDSYGNV-ARLICVIDGDEDDMRMAVYKAIYND-----NIKKP----- 152

QY 118 VGAGAYREVEADPGRLAVEALVTRRCVCVAQRWGGKREVMYTAFK--ALGDSVDYVQV 175
Db 153 -----EFVLCSDDKEGERIDS--DFSRDICVLQPHRGKRECLYTGFLAKMDPSVNAVVL 206

QY 176 CDSDFRLDPMLLELVRLVLDDEPRVGAAGGVDRILNPLDSWVSFLSSLRYVAFNVERAC 235
Db 207 IDSDTVLEKDAILEVYVYPLACDPEIQAVAGECKIWN-TDTLLSLVAVWRYYSAFCVERSA 265

QY 236 QSYFHCVCISGFLGYRNLLQGFLEAWYNOKFLGTHCTFGDDRHLNRLSMG----Y 291
Db 266 QSFRTVQCVGGLKAYKIDIIKEIKDPWISORFLOQKCTYGGDRRLNEILMRGKKVVF 325

QY 292 ATKAEGRWSGTP 304
Db 326 TPFVAVG--WSDSP 336

Search completed: March 11, 2005, 14:44:39
Job time : 72.8182 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 02:13:47 ; Search time 94.4613 Seconds
(without alignments)
10259.279 Million cell updates/sec

Title: US-10-672-399-9
Perfect score: 20
Sequence: 1 cgggcttgcagactactt 20
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_bai.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ats.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	1581	6	CQ731079 Sequence
2	20	100.0	2088	9	HSU59269 Human hyalu
3	20	100.0	2108	9	HUMHAS
4	20	100.0	2109	9	AY463695 Papio anu
5	20	100.0	2116	6	AR220003 Sequence
6	20	100.0	2117	6	AR137038 Sequence
7	20	100.0	2119	9	BC035837 Homo sapi
8	20	100.0	185623	2	AC137058
9	20	100.0	193986	2	AC130783 Pan trogl
10	20	100.0	229155	9	AC018755 Homo sapi
11	18.4	92.0	30	6	E28458 Hyaluronate
12	18.4	92.0	517	4	AB017803
13	18.4	92.0	605	11	BV161384
14	18.4	92.0	621	11	BV099322
15	18.4	92.0	621	11	BV160507
16	18.4	92.0	2095	10	AB097568
17	18.4	92.0	2102	6	E13681 DNA encodin
18	18.4	92.0	2102	6	E28454 Hyaluronate
19	18.4	92.0	2102	6	E30971 Hyaluronc

20	18.4	92.0	2102	6	E34326
21	18.4	92.0	2102	10	MUSHAS
c 22	18.4	92.0	72955	2	AC108651 3
23	18.4	92.0	257325	2	AC116203
c 24	18.4	92.0	293184	2	AC079487
c 25	18	90.0	285648	2	AC129766
c 26	17.4	87.0	1009	5	CR406688
c 27	17.4	87.0	2006	5	DRU31079
28	17.4	87.0	2060	5	BC071301
29	17.4	87.0	91871	9	AY340073
c 30	17.4	87.0	149872	2	AC151565
31	17.4	87.0	170962	2	AL389927
32	17.4	87.0	175594	9	AL162503
c 33	16.8	84.0	562	11	G75028
c 34	16.8	84.0	642	11	BV051323
35	16.8	84.0	938	11	G75183
c 36	16.8	84.0	30382	3	CEP42811
c 37	16.8	84.0	78835	9	AL390754
c 38	16.8	84.0	84701	2	AC016122
c 39	16.8	84.0	111486	2	AC136894
40	16.8	84.0	140491	8	AP003276
41	16.8	84.0	142680	8	AP004611
c 42	16.8	84.0	199050	2	AC102844
c 43	16.8	84.0	215046	2	AF354168
c 44	16.8	84.0	218800	2	AC094410
c 45	16.8	84.0	226362	2	AC094851

ALIGNMENTS

RESULT 1
CQ731079
LOCUS CQ731079 1581 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 17013 from Patent WO02068579.
ACCESSION CQ731079
VERSION CQ731079.1 GI:42306695
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 17013 06-SEP-2002;
FEATURES
source PE Corporation (US)
Location/Qualifiers
1..1581
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN	Query Match	100.0%;	Score 20;	DB 6;	Length 1581;
	Best Local Similarity	100.0%;	Pred. No. 2.6;	Mismatches	0;
	Matches	20;	Conservative	0;	Indels
					0;
Qy	1	CGGGCTTGTCTCAGAGCTACTT	20		
Db	727	CGGGCTTGTCTCAGAGCTACTT	746		
RESULT 2					
HSU59269					
LOCUS	HSU59269	2088 bp	mRNA	linear	PRI 24-SEP-1996
DEFINITION	Human hyaluronan synthase mRNA, complete cds.				
ACCESSION	U59269				
VERSION	U59269.1	GI:1556464			
KEYWORDS					
SOURCE	Homo sapiens (human)				

ORGANISM	Homo sapiens	source	1. .2108
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		/organism="Homo sapiens"
AUTHORS	Shyjan,A.M., Haldin,P., Butcher,E.C., Yoshino,T. and Briskin,M.J.		/mol_type="mRNA"
TITLE	Functional cloning of the cDNA for a human hyaluronan synthase		/db_xref="taxon:9606"
JOURNAL	J. Biol. Chem. 271 (38), 23395-23399 (1996)		/tissue_type="brain"
MEDLINE	96394438	gene	/clone_lib="/lgtil"
PUBMED	8798544		/dev_stages="fetal"
REFERENCE	2 (bases 1 to 2088)		1. .2108
AUTHORS	Briskin,M.J. and Shyjan,A.M.	CDS	/gene="hHAS"
TITLE	Direct Submission		149. .1780
JOURNAL	Submitted (24-MAY-1996) LeukoSite Inc., 215 First Street, Cambridge, MA 02142, USA		/gene="hHAS"
FEATURES	Location/Qualifiers		/codon_start=1
source	1. .2088		/product="hyaluronan synthase"
	/organism="Homo sapiens"		/protein_id="BAAL2351.1"
	/mol_type="mRNA"		/db_xref="GI:1401034"
	/db_xref="taxon:9606"		/translation="MTWAYAGVPLASDRYGLAFGLYGAFLSAHLVAQSLFAYLEHRRVAAARGLDAAATARSVALTISAYQEDPAYLRQCLASARALLYPRARVRVIMVDGNRAEDLYNVDMEFVADDPATYVMDGNYPHQPWEPAAAGVAGAGYREVEADPGLRAVEALVTRRCVCVCAORMGKREVMYTAFAKAGDSVDYVQVCDSDTRLDPMALLLRLVLEDEPRVAGGVGDRILNPLDSVSLSLRYVAFNVERACQSYFHCVSCISGPLGLYRNLLQGFLEAWYKQFLGTHCTFGDDRHLTNRLMSGYATKTSRSRCVSETPSSFLRMLSQOTRSKSYFREWLYNALMWHRRHAWMTYEAUVSLGFFPFVAATVLRIFYAGRPMALLWVLLCVQVALAKAAFAAWLRGCLRMVLLSYAPLYMCGLLPAKFLALVTMNO SGWGTSGRRKLAANYVPLPLALWALLLGLVRSVAHEARADWSGFSRAAEAYHLAA GAGAYGVYVWAMLTLYWVGVRRLCRRRTGGYRVQV"
	ORIGIN		
	Query Match		100.0%; Score 20; DB 9; Length 2108;
	Best Local Similarity		100.0%; Pred. No. 2.6;
	Matches		20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	CGGGCTTGTACAGCTACTT 20	
Db	920	CGGGCTTGTACAGCTACTT 939	
	RESULT 4		
AY463695	LOCUS	AY463695	2109 bp mRNA linear PRI 07-DEC-2003
DEFINITION	Papio anubis hyaluronan synthase (HAS1) mRNA, complete cds.		
ACCESSION	AY463695		
VERSION	AY463695.1	GI:38607341	
KEYWORDS	Papio anubis (olive baboon)		
SOURCE	Papio anubis		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Papio.		
REFERENCE	1 (bases 1 to 2109)		
AUTHORS	Martinez-Duncker,I., Oriol,R. and Mollicone,R.		
TITLE	Evolution of the the hyaluronan, nodulation c, chitin and cellulose synthases: a superfamily of cell-wall associated carbohydrate polymerizing enzymes		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2109)		
AUTHORS	Martinez-Duncker,I., Oriol,R. and Mollicone,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-NOV-2003) Unite 504 Glycobiology Et Signalisation Cellulaire, Inserm, 16 Avenue Paul Vaillant Couturier, Villejuif 94807, France		
FEATURES	Location/Qualifiers		
source	1. .2109		
	/organism="Papio anubis"		
	/mol_type="mRNA"		
	/db_xref="taxon:9555"		
	1. .2109		
	/gene="HAS1"		
	36. .1787		
	/gene="HAS1"		
	/note="cell wall associated protein; produces hyalurononic acid chains"		
	/codon_start=1		
	/product="hyaluronan synthase"		
	ORIGIN		
	Query Match		100.0%; Score 20; DB 9; Length 2088;
	Best Local Similarity		100.0%; Pred. No. 2.6;
	Matches		20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	CGGGCTTGTACAGCTACTT 20	
Db	912	CGGGCTTGTACAGCTACTT 931	
	RESULT 3		
HUMHAS	LOCUS	HUMHAS	2108 bp mRNA linear PRI 06-FEB-1999
DEFINITION	Homo sapiens mRNA for hyaluronan synthase, complete cds.		
ACCESSION	D84424		
VERSION	D84424.1	GI:1401033	
KEYWORDS	hHAS; hyaluronan synthase.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 2108)		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Itano,N. and Kimata,K.		
JOURNAL	Molecular cloning of human hyaluronan synthase		
MEDLINE	96244584		
PUBMED	8651928		
REFERENCE	2 (bases 1 to 2108)		
AUTHORS	Itano,N.		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 2108)		
AUTHORS	Itano,N.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-APR-1996) Naoki Itano, Aichi Medical University, Institute for Molecular Science of Medicine; Nagakute, Aichi-gun, Aichi 480-11, Japan (Tel:052-264-4811(ex.2087), Fax:0561-63-3532)		
FEATURES	Location/Qualifiers		

/protein_id="AAR25554.1"
/db_xref="GI:38607342"
/translation="MTQDTPKTPFAARRCSGLARRVLTIAFALLILGLMTWAYAAGV
PLASRYGLLAFGLYGAFLSAHLLAQSLPAYLEHRRVAAARAAARGRLDAATARSV
ALTISAYQEDPAYLRQCILVSARALLYPRARLVLMVDGNRPEDLIYNDMREVPFAD
DPATYVNDGNTHQPEPAAGVAGVATREVEAEDPGRILAVALVTRRCVCAQRWG
GKREVYTFKALGSDVDYQVCDSDTRLDPMALDELVLQVLDDEPRGAVGVDVRLN
PLDSVYFSLSLRYVAFNVERACQSYFHCYSCISGPLGLYNNILQQFLFLEAWYKQF
LGTHCTGDDRHLLTNRLMSMGVATKYTSRSCYSETSPSFLRWLSQQTRWSKSYFREW
LYNALWHRHHAWMTYEAVVSGLPFPFVAATVLRFLFYAGRPWALLWLLVCQVALAK
RAFAARGLCEMVLILSLYAPLYMCGLLPAKFLALVYTNQSGWGTSGRRKLAANYPL
LPLALWALLLGLVRSVAHEARADWSGFSRAEYHLAAGAGATVGYVWVWLLTLYWV
GVRRLCRRTTGGYRVQV"

ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 2109;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACAGCTACTT 20
|||||
Db 927 CGGGCTTGTGACAGCTACTT 946

RESULT 5
AR220003 LOCUS 2116 bp mRNA linear PAT 26-SEP-2002
DEFINITION Sequence 1 from patent US 6423514.
ACCESSION AR220003
VERSION AR220003.1 GI:23324403

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE
1 (bases 1 to 2116)
AUTHORS Briskin,M.J.

TITLE Mammalian hyaluronan synthases, nucleic acids and uses thereof
JOURNAL Patent: US 6423514-A 1 23-JUL-2002;
FEATURES Location/Qualifiers

1..2116
source
/organism="unknown"
/mol_type="mRNA"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2116;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACAGCTACTT 20
|||||
Db 912 CGGGCTTGTGACAGCTACTT 931

RESULT 6
AR137038 LOCUS 2117 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6162908.
ACCESSION AR137038
VERSION AR137038.1 GI:14478288

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE
1 (bases 1 to 2117)

AUTHORS Itano,N. and Kimata,K.

TITLE Polypeptide of human-origin hyaluronate synthetase and DNA encoding
the same

JOURNAL Patent: US 6162908-A 1 19-DEC-2000;

FEATURES Location/Qualifiers

1..2117
source

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2117;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACAGCTACTT 20
|||||
Db 920 CGGGCTTGTGACAGCTACTT 939

RESULT 7
BC035837 LOCUS 2119 bp mRNA linear PRI 30-JUN-2004

DEFINITION Homo sapiens hyaluronan synthase 1, mRNA (cDNA clone MGC:46218
IMAGE:5589083), complete cds.

ACCESSION BC035837
VERSION BC035837.1 GI:23243101

KEYWORDS
SOURCE MGC.

ORGANISM Homo sapiens (human)

REFERENCE
1 (bases 1 to 2119)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Maman,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shvachenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinaki,M.I., Skalek,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE human and mouse cDNA sequences

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2119)

Strausberg,R.

Direct Submission

Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@hgrl.nih.gov

Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgonov,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,J.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 79 Row: f Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504338.

FEATURES
source
1..2119
/organism="Homo sapiens"
/no_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:46218 IMAGE:5589083"
/tissue="Ovary, pooled from 3 adults"
/clone_lib="NIH MGC_125"
/lab_host="DH10B"
/note="Vector: PCMV-SPORE6"
1..2119
/gene="HAS1"
/note="synonym: HAS"
/db_xref="LocusID:3036"
/db_xref="MIM:601463"
42..1775
/gene="HAS1"
/codon_start=1
/product="hyaluronan synthase 1"
/protein_id="AAH35837.1"
/db_xref="GI:23243102"
/db_xref="LocusID:3036"
/db_xref="MIM:601463"
/translation="MRQAPKPTPAARCSGLARRVLTAFALILGLMTWYAAAGVP
LQSDYGLFGLGALVAHLVAQSLPAYLEHRRVAAARGLDAAATARSVALTISA
YODYPAYLQCLASARALLYPRLRLVMDVGNRAEDLVMDVDFVAFDEDPATVY
WDGNYHPWEPAAAGVAGAGAYREAEADPCGLAVEALVTRCVCVAORWGKREVM
YTFKALGDSDHYQVQSDTRLDPMALLELVRLVLEDEPRVAGVGGDVRIINPLDSWV
YFLSLRLVAPNVERACQSYFHCVCISGLPLGYRLNLQQLFLEAWYKQFLGTHCT
FGDDHLTRNLSMGYATKYTSRSRCYSETSPSFLWLSQOTRWSKSYFREWLYNALW
WRHHAWYTYEAVSGLPFFVFAATVLRFLFYAGRPWALLVLLCVQVALAKAAPAAW
LGLGLRWLLSLYAPLYWCGLLPAKFLALVTMNGSGCTSGRRKLAANYVPLPLALW
ALLGLGLVRSVHAERDMSGPSRAAAYHLAAGAGVGVYWMVLTLYWGVRRLC
RRRTGGYRQV"

gene

CDS

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 2119;
Best Local Similarity 100.0%; Pred. No. 2.6; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0

Qy 1 CGGGCTTGTACAGGACTACTT 20
|||||
Db 915 CGGGCTTGTACAGGACTACTT 934
|||||

RESULT 8

AC137058/c
LOCUS
DEFINITION AC137058 185623 bp DNA linear HTG 21-FEB-2003
Papio anubis clone RP41-126M5, WORKING DRAFT SEQUENCE, 10 ordered
pieces.
ACCESSION AC137058
VERSION AC137058.3 GI:28460766
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Papio anubis (olive baboon)
ORGANISM Papio anubis

REFERENCE

AUTHORS
Akter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Mastello,C., Maskeri,B., McDowell,J.,
Paquirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 185623)
Green,E.D.
Direct Submission
Submitted (15-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 185623)
Green,E.D.
Direct Submission
Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Feb 21, 2003 this sequence version replaced gi:27476124.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: dtj
Center clone name: 126M05

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 183052 bases at least Q40
Consensus quality: 183851 bases at least Q30
Consensus quality: 184434 bases at least Q20
Insert size: 160000; agarose-fp
Insert size: 184723; sum-of-contigs
Quality coverage: 14.13x in Q20 bases; agarose-fp
Quality coverage: 12.24x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 9984: contig of 9984 bp in length
* 9985 10084: gap of unknown length
* 10085 49805: contig of 39721 bp in length
* 49806 49905: gap of unknown length
* 49906 66936: contig of 17031 bp in length
* 66937 67036: gap of unknown length
* 67037 83604: contig of 16568 bp in length
* 83704 83705 108246: gap of unknown length
* 108246 108346: contig of 24542 bp in length
* 108347 112021: contig of 3675 bp in length
* 112021 112121: gap of unknown length
* 112122 134817: contig of 22696 bp in length
* 134818 134917: gap of unknown length
* 134918 147290: contig of 12373 bp in length
* 147291 147390: gap of unknown length
* 147391 185337: contig of 37947 bp in length
* 185338 185437: gap of unknown length


```

FEATURES
  * 185438 185623: contig of 186 bp in length.
    Location/Qualifiers
      1..185623
        /organism="Papio anubis"
        /mol_type="genomic DNA"
        /db_xref="taxon:9555"
        /clone="RP41-126M5"
        /clone_lib="RP41"
      1..15263
        /note="clone overlaps with GenBank Accession Number
        AC130273 clone RP41-30706 (center project name dsu)"
      1..9984
        /note="assembly_fragment
        clone_end:T7
        vector_side:left"
      10085..49805
        /note="assembly_fragment"
      49906..66936
        /note="assembly_fragment"
      67037..83604
        /note="assembly_fragment"
      83705..108246
        /note="assembly_fragment"
      108347..112021
        /note="assembly_fragment"
      112122..134817
        /note="assembly_fragment"
      134918..147290
        /note="assembly_fragment"
      147391..185337
        /note="assembly_fragment"
      185438..185623
        /note="assembly_fragment
        clone_end:SP6
        vector_side:right"

ORIGIN
  Query Match 100.0%; Score 20; DB 2; Length 185623;
  Best Local Similarity 100.0%; Pred. No. 3.4;
  Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTCTCAGAGTACTT 20
    |||||
Db 135015 CGGGCTTCTCAGAGTACTT 134996

RESULT 9
AC130783/c AC130783 193986 bp DNA linear HTG 19-NOV-2002
LOCUS Pan troglodytes clone CH251-426A12, WORKING DRAFT SEQUENCE, 8
DEFINITION ordered pieces.
ACCESSION AC130783
VERSION AC130783.2 GI:25100968
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 193986)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
AUTHORS Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaapi,R., Maduro,Q.L., Maduro,V.B.,
Marques,E.H., Mastello,C., Maskeri,B., McDowell,J.,
Pacurigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
UNPUBLISHED
TITLE NISC
JOURNAL NISC
REFERENCE 2 (bases 1 to 193986)
AUTHORS Green,E.D.

```

```

TITLE Direct Submission
JOURNAL Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Grovetmont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 193986)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovetmont Circle, Gaithersburg, MD 20877, USA
COMMENT On Nov 19, 2002 this sequence version replaced gi:22218453.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: dxid
Center clone name: 426A12

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 151764 bases at least Q40
Consensus quality: 192510 bases at least Q30
Consensus quality: 192998 bases at least Q20
Insert size: 184000; agarose-fp
Quality coverage: 11.49x in Q20 bases; agarose-fp
Quality coverage: 10.94x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 11512: contig of 11512 bp in length
* 11513 11612: gap of unknown length
* 11613 59488: contig of 47876 bp in length
* 59489 59588: gap of unknown length
* 59589 103266: contig of 43678 bp in length
* 103267 103366: gap of unknown length
* 103367 127885: contig of 24519 bp in length
* 127886 127985: gap of unknown length
* 127986 129692: contig of 1707 bp in length
* 129693 129792: gap of unknown length
* 129793 131077: contig of 1285 bp in length
* 131078 131177: gap of unknown length
* 131178 174663: contig of 43486 bp in length
* 174664 174763: gap of unknown length
* 174764 193986: contig of 19223 bp in length.
FEATURES
  1..193986
    /organism="Pan troglodytes"
    /mol_type="genomic DNA"
    /db_xref="taxon:9598"
    /clone="CH251-426A12"
    /clone_lib="CH251"
  source

```

```

misc_feature      1. .11512
/note="assembly_fragment
clone_end:SP6
vector_side:left"
1. .76
/note="clone overlaps with GenBank Accession Number
AC132069 clone CH251-132A6 (center project name dxc)"
11613. .59488
/note="assembly_fragment"
59589. .103266
/note="assembly_fragment"
103367. .127885
/note="assembly_fragment"
127986. .129692
/note="assembly_fragment"
129793. .131077
/note="assembly_fragment"
131178. .174663
/note="assembly_fragment"
174764. .193986
/note="assembly_fragment
clone_end:T7
vector_side:right"

ORIGIN
Query Match      100.0%; Score 20; DB 2; Length 193986;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTAGAGCTACTT 20
|||||
Db 129885 CGGGCTTGTCTAGAGCTACTT 129866

RESULT 10
AC018755/c
LOCUS      AC018755      229155 bp      DNA      linear      PRI 26-JUL-2000
DEFINITION Homo sapiens chromosome 19, BAC BC330783 (CIT-HSPC_470E3), complete
sequence.
ACCESSION  AC018755
VERSION    AC018755.3      GI:9454515
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

REFERENCE
AUTHORS    Lamerdin,J.E., McCready,P.M., Skowronski,E., Viswanathan,V.,
1 (bases 1 to 229155)
Burkhardt-Schultz,K., Gordon,L., Dias,J., Ramirez,M., Stilwagen,S.,
Phan,H., Velasco,N., Do,L., Regala,W., Terry,A., Brower,A.,
Garnes,J., Danganan,L., Erler,A., Christensen,M., Georgescu,A.,
Avila,J., Liu,S., Andreise,T., Frankheim,M., Attix,C.,
Amico-Keller,G., Cosfield,J., Duarte,S., Lucas,S., Bruce,R.,
Thomas,P., Quan,G., Kronmiller,B., Arellano,A., Sanders,C., Ow,D.,
Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and Carrano,A.V.
Sequence analysis of a 5-Mb region in 19q13.4

JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 229155)
AUTHORS    DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL    Submitted (19-DEC-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 229155)
AUTHORS    Lamerdin,J.E.
TITLE      Direct Submission
JOURNAL    Submitted (26-JUL-2000) DOE Joint Genome Institute, Lawrence
Livermore National Laboratory, Livermore, CA 94550
COMMENT    On Jul 26, 2000 this sequence version replaced gi:7458780.
Map and sequence oriented from centromere to q-telomere. BC330783
(CIT-HSPC_470E3) is overlapped on the left by BC849408
and overlaps cosmid R28782 (L1NL-R 248F10, AC005946) on the right
from bases 217,905 to 229,155 of this accession. Additional

chromosome 19 map and sequence information may be obtained at:
http://www-bio.llnl.gov/bhrp/genome/genome.html.
FEATURES
source
1. .229155
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
/clone="CTC-470E3 (L1NL clone name BC330783)"
/clone_lib="CalTech human BAC library C"
201. .262
/rpt_family="L2"
repeat_region
<299. .16870
gene
/genes="ZNF175"
/note="zinc finger protein 175; synonym: OTK18"
join(<299. .550,8540. .8666,8936. .9031,13776. .16870)
/mRNA
/genes="ZNF175"
/product="ZNF175 mRNA"
CDS
join(479. .550,8540. .8666,8936. .9031,13776. .15616)
/genes="ZNF175"
/codon_start=1
/product="zinc finger protein 175"
/protein_id="AAF87843.1"
/db_xref="GI:9454517"
/translation="MPADVNLKQKQVLPGEKQDGCSEASVSFEDVTVDPSREEWQOL
DPQRCHYRDVLMLEYSHLFVAGVTHIPNPFVIFMLKEKEPRVEAEVSHQRCQEREF
GLIEPQKIEKKKASQKDMVGFTRDGSWCSIIIEELRLDADRTKKDQONQIQPWSHSA
FTEKTLNTNSECYKDPQGMVIRTRPHLASQKQKCCLPFTESKLNLNLYNQNESN
DFEQLDDVSGSGYFQSHSSSDACSKNIHTGETFCCKGCRKVCCHKQSLQKHQHTQK
KPDGSCGSGSFTQKSHLFAQORIHVSVGNLHEGCKGKAFMPQLKLSVLTHTGDIIP
CICCEGKGFQIQRSELLTHQKTHTRKPKYCHDGGKAFQFQMSLFRHQRTSREKLYE
CSCEGFGFSONSLIIHOKIHTGRQYACSECCGKAFQKSTLSLQHIHSGQKSYVCI
EGCOAFQKAHLIVHQRSHTGKPYQCHNCGKSFISKQLDIHRIHTGKPYECSDC
KCTFTQKSHLNIHQHTGKPYHVCCEGKAFNQKSLSMHQRIHTGKPYKCECGK
AFTSKSQFKEHQRIHTGKPYVCTGCEGKAFNGRSNFHKOIHTHTRPFPVYKCGKAF
VQKSELTHQTHMGKPYECLDQKGSFKKPKLKVHQRIHTGKPYVCEGKAFNFN
RSNPNKHQTHTRDQSKYSYKGFQK"
1097. .1198
/rpt_family="MER99"
repeat_region
1559. .1829
/rpt_family="AluJb"
repeat_region
complement(2130. .2355)
/rpt_family="L1ME"
repeat_region
complement(2359. .2627)
/rpt_family="AluY"
repeat_region
complement(2816. .3183)
/rpt_family="HAL1"
repeat_region
complement(3290. .3598)
/rpt_family="AluSg"
repeat_region
complement(3611. .3760)
/rpt_family="MIR"
repeat_region
complement(3766. .4067)
/rpt_family="HAL1"
repeat_region
4098. .4215
/rpt_family="AluJb"
repeat_region
complement(4339. .4406)
/rpt_family="HAL1"
repeat_region
complement(4518. .4724)
/rpt_family="HAL1"
repeat_region
complement(4803. .5110)
/rpt_family="AluY"
repeat_region
5286. .5561
/rpt_family="AluSg"
repeat_region
5635. .5932
/rpt_family="AluJo"
repeat_region
5984. .6120
/rpt_family="AluJb"
repeat_region
6121. .6423
/rpt_family="AluSx"
repeat_region
6424. .6579
/rpt_family="AluJb"
repeat_region
6647. .6913

```

```
repeat_region      /rpt_family="L1ME"
6936. .7295
/rpt_family="MER33"
7700. .7785
/rpt_family="L2"
complement(9216. .9434)
/rpt_family="L1ME"
complement(9435. .9745)
/rpt_family="AluY"
complement(9746. .9815)
/rpt_family="L1ME"
complement(10009. .10303)
/rpt_family="AluSp"
11783. .12637
/rpt_family="L1MB2"
12180. .12330
/gene="ZNF175"
/standard_name="SHGC-32377"
13436. .13619
/rpt_family="MIR"
16068. .16112
/rpt_family="(TA)n"
complement(16113. .16329)
/rpt_family="MER46A"
16330. .16390
/rpt_family="(TA)n"
16403. .16439
/rpt_family="(TA)n"
16793. .17013
/rpt_family="(TATG)n"
/rpt_family="MER58A"
17024. .17052
/rpt_family="(CA)n"
17311. .17545
/rpt_family="MIR"
complement(18683. .18896)
/rpt_family="MIR"
complement(18915. .19183)
19336. .19599
/rpt_family="AluY"
/rpt_family="MIR"
19811. .19928
/rpt_family="GA-rich"
22181. .23081
/rpt_family="L2"
complement(23692. .23901)
/rpt_family="L2"
complement(23902. .24201)
/rpt_family="AluSg1"
complement(24202. .24651)
/rpt_family="L2"
26136. .26434
/rpt_family="LTR12"
26435. .26525
/rpt_family="LTR30"
26526. .26643
/rpt_family="LTR12"
complement(26686. .26884)
/rpt_family="AluSg/x"
27065. .27240
/rpt_family="MIR"
complement(27238. .27554)
/rpt_family="L2"
28406. .28433
/rpt_family="GC-rich"
28641. .28678
/rpt_family="GC-rich"
complement(28997. .29259)
/rpt_family="MSR1"
complement(30219. .30489)
/rpt_family="L2"
30753. .30882
/rpt_family="FLAM_C"
31138. .31416

repeat_region      /rpt_family="AluY"
complement(31547. .31716)
/rpt_family="MER3"
32041. .32112
/rpt_family="L1MBc"
32117. .32516
/rpt_family="MLT2E"
35056. .35118
/rpt_family="MIR"
complement(35578. .35636)
/rpt_family="MER5A"
complement(35690. .35820)
/rpt_family="MIR"
35872. .35902
/rpt_family="(TG)n"
complement(35906. .36209)
/rpt_family="AluJo"

Query Match      100.0%; Score 20; DB 9; Length 229155;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches: 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGGGCTTGTCTCAGAGCTACTT 20
Db      144168 CGGGCTTGTCTCAGAGCTACTT 144149

RESULT 11
E28458 LOCUS
E28458 Hyaluronate synthase promoter DNA.
DEFINITION
ACCESSION E28458
VERSION E28458.1 GI:13018350
KEYWORDS JP 1999196875-A/6.
SOURCE
ORGANISM
unclassified
unclassified
REFERENCE 1 (bases 1 to 30)
AUTHORS Voichi, Y., Naoki, I. and Koji, K.
TITLE Hyaluronate synthase promoter DNA
JOURNAL Patent: JP 1999196875-A 6 27-JUL-1999;
SEIRAGAKU KOGYO CO LTD
COMMENT OS Unidentified
PN JP 1999196875-A/6
PD 27-JUL-1999
PF 14-JAN-1998 JP 1998006191
PR YOICHI YAMADA, NAOKI ITANO, KOJI KIMATA
PI C12N15/09, C12N9/00, C12Q1/68//C12N15/09, C12R1:91, C12N15/00,
PC (C12N15/00, C12R1:91)
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1. .30
FT Location/Qualifiers
FT 1. .30
FT /organism="Unidentified".
FT /organism="unidentified"
FT /mol_type="genomic DNA"
FT /db_xref="taxon:32644"

FEATURES
source
1. .30
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match      92.0%; Score 18.4; DB 6; Length 30;
Best Local Similarity 95.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CGGGCTTGTCTCAGAGCTACTT 20
Db      7 CGAGCTTGTCTCAGAGCTACTT 26

RESULT 12
AB017803
```

LOCUS AB017803 517 bp mRNA linear MAM 02-APR-1999
 DEFINITION Bos taurus mRNA for hyaluronan synthase 1, partial cds.
 ACCESSION AB017803
 VERSION AB017803.1 GI:4586933
 KEYWORDS HAS1; hyaluronan synthase 1.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.
 REFERENCE 1 (bases 1 to 517)
 AUTHORS Yamashita,H., Usui,T. and Suzuki,K.
 TITLE HAS1
 JOURNAL Published Only in Database (1999)
 REFERENCE 2 (bases 1 to 517)
 AUTHORS Yamashita,H., Usui,T. and Suzuki,K.
 TITLE Direct Submission
 JOURNAL Submitted (21-SEP-1998) Hidetoshi Yamashita, Faculty of Medicine,
 University of Tokyo, Department of Ophthalmology; Hongo 7-3-1,
 Bunkyo-ku, Tokyo 113-8655, Japan
 (E-mail:hyama-tky@umin.u-tokyo.ac.jp, Tel: +81-3-3815-5411 (ex.3494),
 Fax: +81-3-3817-0798)

FEATURES

source

1..517
 Location/Qualifiers
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 1..517
 /gene="HAS1"
 <1..>517
 /gene="HAS1"
 /codon_start=1
 /product="hyaluronan synthase 1"
 /protein_id="BAA76546.1"
 /db_xref="GI:4586934"
 /translation="RACOSYFHCVCISGPIGLYRNLLQOLFLEAWYKQELGTHCTF
 GDDRHLTRMLSMGVATKYTSRSCYSETPSSFLWLSQOTRWKSYFREWLYNALW
 HRHWAMTYEAVSGLFFPFFVNAIVLRIFYAGRPWALLWLLVCQGVALLAKAAPAWL
 RGLRMVLVLSLY"

gene

CDS

ORIGIN

Query Match 92.0%; Score 18.4; DB 4; Length 517;
 Best Local Similarity 95.0%; Pred. No. 22;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCAGAGTACTT 20

Db 1 CGGGCTTGCCAGAGTACTT 20

RESULT 13

BV161384

LOCUS

DEFINITION BV161384 605 bp DNA linear STS 15-MAY-2004
 tagged site.

ACCESSION

BV161384

VERSION

BV161384.1

KEYWORDS

STS.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 605)
 REFERENCE 1 (bases 1 to 605)
 AUTHORS Usuka,J., Liao,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,
 McPherson,J.D., Foerzler,D. and Peltz,G.
 TITLE Mus musculus SNPs
 JOURNAL Unpublished (2003)
 COMMENT

Contact: Jonathan Usuka

Roche Palo Alto Genetics and Genomics Department

Roche Palo Alto

3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA

Tel: 6508555807

FEATURES

source

1..605
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /map="17-12488-11884-CAAA01219574.1.1.20118"
 /clone_lib="Roche Palo Alto"
 /note="SNPs developed from assay sequences derived from 15
 different strains of mice (as of October 1, 2003). Those
 strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/cSnJ,
 BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J,
 MRL/MpJ, NZB/BinJ, NZW/LaC, SPRET/Ei."
 <1..>605

STS

ORIGIN

Query Match 92.0%; Score 18.4; DB 11; Length 605;
 Best Local Similarity 95.0%; Pred. No. 22;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCAGAGTACTT 20

Db 440 CGAGCTTGTCAGAGTACTT 459

RESULT 14

BV099322/c

LOCUS

DEFINITION BV099322 621 bp DNA linear STS 14-JAN-2004
 tagged site.

ACCESSION

BV099322

VERSION

BV099322.1

KEYWORDS

STS.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 621)
 REFERENCE 1 (bases 1 to 621)
 AUTHORS Usuka,J., Liao,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,
 McPherson,J.D., Foerzler,D. and Peltz,G.
 TITLE Mus musculus SNPs
 JOURNAL Unpublished (2003)
 COMMENT

Contact: Jonathan Usuka

Roche Palo Alto Genetics and Genomics Department

Roche Palo Alto

3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA

Tel: 6508555807

Email: Jonathan.Usuka@roche.com

Primer A: No primer submitted with this STS

Primer B: No primer submitted with this STS.

FEATURES

source

1..621
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /map="17-11968-12587-CAAA01219574.1.1.20118"
 /clone_lib="Roche Palo Alto"
 /note="SNPs developed from assay sequences derived from 15
 different strains of mice (as of October 1, 2003). Those
 strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/cSnJ,
 BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J,
 MRL/MpJ, NZB/BinJ, NZW/LaC, SPRET/Ei."
 <1..>621

STS

ORIGIN

Query Match 92.0%; Score 18.4; DB 11; Length 621;
 Best Local Similarity 95.0%; Pred. No. 22;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCAGAGTACTT 20

Db 82 CGAGCTTGTTCAGAGTACTT 63

RESULT 15
BV160507/c

LOCUS BV160507 621 bp DNA linear STS 15-MAY-2004

DEFINITION RPAMMSQ0038182 Roche Palo Alto Mus musculus STS genomic, sequence tagged site.

ACCESSION BV160507

VERSION BV160507.1 GI:47263410

KEYWORDS STS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 621)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Usuka,J., Liao,G., Cheng,J., Nguyen,A., Bach,C., Puech,A., McPherson,J.D., Foernzler,D. and Feltz,G.

TITLE Mus musculus SNPs

JOURNAL Unpublished (2003)

COMMENT

Contact: Jonathan Usuka
Roche Palo Alto Genetics and Genomics Department
Roche Palo Alto
3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
Tel: 6508555807
Email: Jonathan.Usuka@roche.com
Primer A: No primer submitted with this STS
Primer B: No primer submitted with this STS.

FEATURES
source
1..621
Location/Qualifiers
1..621
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/map="17-11968-12587-CAAA01219574.1.1.20118"
/clone_lib="Roche Palo Alto"
/note="SNPs developed from assay sequences derived from 15 different strains of mice (as of October 1, 2003). Those strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/cSnJ, BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J, MRL/MpJ, NZB/BinJ, NZW/Lac, SPRET/Ei."
<1..>621

STS

ORIGIN

Query Match 92.0%; Score 18.4; DB 11; Length 621;
Best Local Similarity 95.0%; Pred.No.22;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCTTGTTCAGAGTACTT 20
|||

Db 82 CGAGCTTGTTCAGAGTACTT 63

Search completed: March 13, 2005, 08:01:38
Job time : 97.4613 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 01:47:21 ; Search time 12.9978 Seconds
(without alignments)
9108.809 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggctgtcagagctactt 20

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	490	ACH39998	ACH39998 Human foe
2	20	100.0	1737	ADL13691	ADL13691 Osteoarth
3	20	100.0	1737	ADL13692	ADL13692 Osteoarth
4	20	100.0	2087	ADL13694	ADL13694 Osteoarth
5	20	100.0	2088	ADL13690	ADL13690 Osteoarth
6	20	100.0	2116	AAT99541	AAT99541 Human hya
7	20	100.0	2117	AAT96713	AAT96713 Human hya
8	20	100.0	231222	ADL13693	ADL13693 Osteoarth
9	18.4	92.0	30	AAT10866	AAT10866 PCR prime
10	18.4	92.0	1752	ABZ76734	ABZ76734 Mouse hya
11	18.4	92.0	2102	AAT91855	AAT91855 Mouse hya
12	18.4	92.0	2102	AAT91855	AAT91855 Mouse hya
13	18.4	92.0	2102	AAZ10862	AAZ10862 Hyaluroa
14	18.4	92.0	2102	AAZ88199	AAZ88199 Mouse hya
15	18.4	92.0	2102	AAZ39987	AAZ39987 Murine HA
16	16.4	82.0	25	AAT10869	AAT10869 PCR prime
17	16.4	82.0	2000	ADA73353	ADA73353 Rice gene
18	16.4	82.0	3931	ADF55687	ADF55687 Mouse car
19	16.4	82.0	21410	ACN44668	ACN44668 Mouse gen
20	15.8	79.0	316	ADP93649	ADP93649 Cotton ex

21	15.8	79.0	400	2	AAV78369	AAV78369 Staphyloc
22	15.8	79.0	430	6	ABN22318	ABN22318 Human ORF
23	15.8	79.0	477	4	AAI11803	AAI11803 Probe #17
24	15.8	79.0	477	4	ABA53497	ABA53497 Human foe
25	15.8	79.0	477	4	AAI33119	AAI33119 Probe #18
26	15.8	79.0	477	4	ABA43081	ABA43081 Human bre
27	15.8	79.0	477	4	ABA23262	ABA23262 Probe #17
28	15.8	79.0	477	4	AAK27223	AAK27223 Human bon
29	15.8	79.0	477	4	AAK01768	AAK01768 Human bra
30	15.8	79.0	477	4	ABS26800	ABS26800 Human liv
31	15.8	79.0	477	5	AAI01737	AAI01737 Probe #17
32	15.8	79.0	477	6	ABS01762	ABS01762 Human gen
33	15.8	79.0	582	4	ABA60297	ABA60297 Human foe
34	15.8	79.0	582	4	AAI40181	AAI40181 Probe #88
35	15.8	79.0	582	4	AAK34458	AAK34458 Human bon
36	15.8	79.0	582	4	AAK08575	AAK08575 Human bra
37	15.8	79.0	582	4	ABS34235	ABS34235 Human liv
38	15.8	79.0	1950	8	ACF74273	ACF74273 Staphyloc
39	15.8	79.0	1953	8	ACA20042	ACA20042 Prokaryot
40	15.8	79.0	2111	2	AAT80389	AAT80389 Staphyloc
41	15.8	79.0	9623	2	AAV74477	AAV74477 Staphyloc
42	15.8	79.0	14781	4	AAK66710	AAK66710 Human imm
43	15.8	79.0	42379	12	ADQ97660	ADQ97660 Mouse can
44	15.8	79.0	50720	13	ABD33468	ABD33468 Murine ca
45	15.8	79.0	77834	11	ACN44076	ACN44076 Mouse gen

ALIGNMENTS

RESULT 1

ACH39998
ID ACH39998 standard; cDNA; 490 BP.

XX ACH39998;

XX DT 13-OCT-2003 (first entry)

XX DE Human foetal brain cDNA #1365.

XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.

XX OS Homo sapiens.

XX PN US2003073623-A1.

XX PD 17-APR-2003.

XX PF 30-JUL-2001; 2001US-00918995.

XX PR 30-JUL-2001; 2001US-00918995.

XX PA (DRMA/) DRMANAC R T.

XX PA (LABA/) LABAT I.

XX PA (STAC/) STACHE-CRAIN B.

XX PA (DICK/) DICKSON M C.

XX PA (JONE/) JONES L W.

XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX DR WPI; 2003-615964/58.

XX PT New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.

XX PS Claim 1; SEQ ID NO 27210; 44pp; English.

XX CC The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also


```
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTTCAGAGCTACTT 20
Db 877 CGGGCTTGTTCAGAGCTACTT 896

RESULT 4
ADL13694
ID ADL13694 standard; DNA; 2087 BP.
XX AC ADL13694;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #226.
XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX DR WPI; 2003-559141/52.
XX PT Determining susceptibility of an individual to joint space narrowing,
XX PT osteophyte development and/or joint pain comprises identifying whether
XX PT the individual has at least one polymorphism in a polynucleotide encoding
XX PS a protein.
XX PS Disclosure; SEQ ID NO 226; 297pp; English.
XX CC The invention relates to a method of determining susceptibility of an
XX CC individual to joint space narrowing and/or osteophyte development and/or
XX CC joint pain comprising identifying whether the individual has at least one
XX CC polymorphism in a polynucleotide encoding at least one of the protein
XX CC listed in the specification. The methods, composition and agent are
XX CC useful for modulating the susceptibility of an individual to joint space
XX CC narrowing and/or osteophyte development and/or joint pain that is
XX CC associated with a disease, preferably osteoarthritis. The cell line and
XX CC the non-human animal are useful for screening for an agent for diagnosing
XX CC an individual having susceptibility to joint space narrowing and/or
XX CC osteophyte development and/or joint pain. This sequence corresponds to
XX CC the polynucleotide encoding a protein listed in the specification. (Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences).
XX SQ Sequence 2087 BP; 302 A; 650 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 2087;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTTCAGAGCTACTT 20
Db 912 CGGGCTTGTTCAGAGCTACTT 931

RESULT 5
ADL13694
ID ADL13694 standard; DNA; 2087 BP.
XX AC ADL13694;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #226.
XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX DR WPI; 2003-559141/52.
XX PT Determining susceptibility of an individual to joint space narrowing,
XX PT osteophyte development and/or joint pain comprises identifying whether
XX PT the individual has at least one polymorphism in a polynucleotide encoding
XX PS a protein.
XX PS Disclosure; SEQ ID NO 226; 297pp; English.
XX CC The invention relates to a method of determining susceptibility of an
XX CC individual to joint space narrowing and/or osteophyte development and/or
XX CC joint pain comprising identifying whether the individual has at least one
XX CC polymorphism in a polynucleotide encoding at least one of the protein
XX CC listed in the specification. The methods, composition and agent are
XX CC useful for modulating the susceptibility of an individual to joint space
XX CC narrowing and/or osteophyte development and/or joint pain that is
XX CC associated with a disease, preferably osteoarthritis. The cell line and
XX CC the non-human animal are useful for screening for an agent for diagnosing
XX CC an individual having susceptibility to joint space narrowing and/or
XX CC osteophyte development and/or joint pain. This sequence corresponds to
XX CC the polynucleotide encoding a protein listed in the specification. (Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences).
XX SQ Sequence 2087 BP; 302 A; 650 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 2087;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTTCAGAGCTACTT 20
Db 912 CGGGCTTGTTCAGAGCTACTT 931

RESULT 6
AAT99541
ID AAT99541 standard; cDNA; 2116 BP.
XX AC AAT99541;
XX DT 21-MAY-1998 (first entry)
XX DE Human hyaluronan synthase cDNA clone 30C.
```

```
ADL13690
ID ADL13690 standard; DNA; 2088 BP.
XX AC ADL13690;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #222.
XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX DR WPI; 2003-559141/52.
XX PT Determining susceptibility of an individual to joint space narrowing,
XX PT osteophyte development and/or joint pain comprises identifying whether
XX PT the individual has at least one polymorphism in a polynucleotide encoding
XX PS a protein.
XX PS Disclosure; SEQ ID NO 222; 297pp; English.
XX CC The invention relates to a method of determining susceptibility of an
XX CC individual to joint space narrowing and/or osteophyte development and/or
XX CC joint pain comprising identifying whether the individual has at least one
XX CC polymorphism in a polynucleotide encoding at least one of the protein
XX CC listed in the specification. The methods, composition and agent are
XX CC useful for modulating the susceptibility of an individual to joint space
XX CC narrowing and/or osteophyte development and/or joint pain that is
XX CC associated with a disease, preferably osteoarthritis. The cell line and
XX CC the non-human animal are useful for screening for an agent for diagnosing
XX CC an individual having susceptibility to joint space narrowing and/or
XX CC osteophyte development and/or joint pain. This sequence corresponds to
XX CC the polynucleotide encoding a protein listed in the specification. (Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences).
XX SQ Sequence 2088 BP; 302 A; 652 C; 711 G; 423 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 2088;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTTCAGAGCTACTT 20
Db 912 CGGGCTTGTTCAGAGCTACTT 931

RESULT 6
AAT99541
ID AAT99541 standard; cDNA; 2116 BP.
XX AC AAT99541;
XX DT 21-MAY-1998 (first entry)
XX DE Human hyaluronan synthase cDNA clone 30C.
```

KW Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
 KW wound healing; vulnery; tissue repair; scar; keloid; therapy; ds.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 36..1772
 FT CDS /tag= a
 FT polyA_signal 2066..2071
 FT /tag= b

XX WO9740174-A1.
 XX 30-OCT-1997.
 XX 17-APR-1997; 97WO-US006350.
 XX 22-APR-1996; 96US-00635552.
 XX (LEUK-) LEUKOSITE INC.
 PA Briskin MJ;
 PI WPI; 1997-549359/50.
 DR P-PSDB; AAW26765.
 XX Human hyaluronan synthase - useful for recombinant production of
 PT hyaluronic acid for wound healing, tissue repair and reducing
 PT hypertrophic scar and keloid formation.
 XX Claim 3; Page 36-38; 58pp; English.

CC cDNA clone 30C includes a coding region for hyaluronan synthase (HAS)
 CC (see AAW26765), an enzyme involved in the synthesis of hyaluronan
 CC (hyaluronic acid) and which has the ability to confer cell adhesion by
 CC the lymphocyte receptor CD44. Clone 30C was isolated using an expression
 CC cloning system developed to isolate cDNA clones that encode proteins that
 CC confer adhesion of the murine T cell lymphoma TK1. A human mesenteric
 CC lymph node expression library was constructed that, upon transfection
 CC into CHO/P cells, yielded clone 30C that mediated rosetting of TK1 cells
 CC to some of the transfectants. The isolated clone can be utilised in a
 CC claimed method for producing HAS in host cells. Such host cells are used
 CC in a claimed method for the production of hyaluronan. Hyaluronan is
 CC useful for wound healing and tissue repair, and can reduce or prevent
 CC hypertrophic scars and keloid formation. It is also used in eye surgery
 CC as a replacement for vitreous fluid

XX Sequence 2116 BP; 330 A; 651 C; 712 G; 423 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 20; DB 2; Length 2116;
 Best Local Similarity 100.0%; Pred. No. 3.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGGGCTTGTCTCAGAGTACTT 20
 |||||
 DB 912 CGGGCTTGTCTCAGAGTACTT 931

RESULT 7
 ID AAT96713
 XX AAT96713 standard; DNA; 2117 BP.
 AC AAT96713;
 XX 22-APR-1998 (first entry)
 DT Human hyaluronate synthetase coding sequence.
 DE Hyaluronate synthetase; human; hyaluronic acid; drug preparation;
 KW cosmetic preparation; gene therapy; carcinogenesis; ss.
 XX Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 149..1780
 FT /tag= a

XX WO9738113-A1.
 XX 16-OCT-1997.
 XX 31-MAR-1997; 97WO-JP001111.
 XX 05-APR-1996; 96JP-00084326.
 XX 30-APR-1996; 96JP-00109663.
 XX (SEK) SEIKAGAKU CORP.
 PA Itano N, Kimata K;
 PI WPI; 1997-512726/47.
 DR P-PSDB; AAW36503.
 XX DNA encoding human hyaluronate synthetase - for industrial scale
 PT production of hyaluronic acid used in generating anti-carcinogenic drugs
 PT or for cosmetics.
 XX Claim 3; Page 23-27; 35pp; Japanese.

CC This sequence encodes a human hyaluronate synthetase, and is the coding
 CC sequence of the invention. The encoded enzyme is useful for industrial
 CC scale production of hyaluronic acid for use in the preparation of drugs
 CC and cosmetics. The drugs can also be used in compositions for the
 CC treatment of disorders involving the lowering of hyaluronic acid
 CC production. The peptides may be used for the preparation of antibodies
 CC recognising hyaluronate synthetase, e.g. for diagnostic purposes.
 CC Antisense DNA or RNA corresponding to the DNA can be used in gene therapy
 CC treatment of carcinogenesis

XX Sequence 2117 BP; 306 A; 673 C; 703 G; 435 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 20; DB 2; Length 2117;
 Best Local Similarity 100.0%; Pred. No. 3.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGGGCTTGTCTCAGAGTACTT 20
 |||||
 DB 920 CGGGCTTGTCTCAGAGTACTT 939

RESULT 8
 ID ADL13693/C
 XX ADL13693 standard; DNA; 231222 BP.
 AC ADL13693;
 XX 06-MAY-2004 (first entry)
 DT Osteoarthritis-associated polymorphic nucleotide #225.
 DE ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
 KW joint space narrowing; osteophyte development; joint pain;
 KW osteoarthritis; SNP; single nucleotide polymorphism.
 OS Homo sapiens.
 XX WO2003054166-A2.
 XX 03-JUL-2003.
 XX 19-DEC-2002; 2002WO-US041225.
 XX 20-DEC-2001; 2001US-0342603P.
 XX (INCY-) INCYTE GENOMICS INC.


```

RESULT 13
AAZ10862
ID AAZ10862 standard; DNA; 2102 BP.
XX
AC AAZ10862;
XX
DT 19-OCT-1999 (first entry)
XX
DE Hyaluronate synthase coding sequence.
XX
KW Hyaluronate synthase; promoter; mouse; cell growth inhibitor; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
CDS 49..1800
FT /*tag= a
FT
XX
PN JP11196875-A.
XX
ED 27-JUL-1999.
XX
PF 14-JAN-1998; 98JP-00006191.
XX
PR 14-JAN-1998; 98JP-00006191.
XX
PA (SEBK ) SEIKAGAKU KOGYO CO LTD.
XX
WPI; 1999-496653/42.
P-PSDB; AAY32503.
DR
DR
PT New promoter DNA of hyaluronate synthase - used to specifically express
PT gene with cell growth inhibiting activity.
XX
PS Example 1; Page 8-10; 13pp; Japanese.
XX
CC This sequence represents the mouse hyaluronate synthase coding sequence,
CC and was used to isolate the hyaluronate synthase promoter of the
CC invention. The promoter can be used for specifically expressing a gene
CC having cell growth inhibiting activity
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;

Query Match 92.0%; Score 18.4; DB 2; Length 2102;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
||| ||||| ||||| ||||| |||||
Db 940 CGAGCTTGTCTCAGAGCTACTT 959

RESULT 14
AAZ88199
ID AAZ88199 standard; cDNA to mRNA; 2102 BP.
XX
AC AAZ88199;
XX
DT 27-APR-2000 (first entry)
XX
DE Mouse hyaluronate synthase modified protein HAS1 encoding cDNA.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
CDS 49..1800
FT /*tag= a
FT /product= "HAS1"
FT /note= "modified hyaluronate synthase protein"
FT
XX
PN JP200004886-A.
XX
PD 11-JAN-2000.
XX
PF 24-JUN-1998; 98JP-00193788.
XX
PR 24-JUN-1998; 98JP-00193788.
XX
PA (SEBK ) SEIKAGAKU KOGYO CO LTD.
XX
WPI; 2000-140125/13.
P-PSDB; AAY68491.
DR
DR
PT A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
PS Claim 18; Page 12-14; 30pp; Japanese.
XX
CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. The present sequence encodes mouse HAS1
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;

Query Match 92.0%; Score 18.4; DB 3; Length 2102;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
||| ||||| ||||| ||||| |||||
Db 940 CGAGCTTGTCTCAGAGCTACTT 959

RESULT 15
AAZ39987
ID AAZ39987 standard; cDNA; 2102 BP.
XX
AC AAZ39987;
XX
DT 19-OCT-2000 (first entry)
XX
DE Murine HAS1 cDNA.
XX
KW Murine; HAS1; hyaluronate synthase; marker gene; intracellular loop;
KW gene targeting; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
CDS 49..1800
FT /*tag= a
FT /product= "HAS1"
FT
XX
PN JP2000116382-A.
XX
PD 25-APR-2000.
XX
PF 13-OCT-1998; 98JP-00291201.
XX
PR 13-OCT-1998; 98JP-00291201.
XX
PA (SEBK ) SEIKAGAKU KOGYO CO LTD.
XX
WPI; 2000-369404/32.
P-PSDB; AAB09948.
DR
DR
PT DNA used for gene targeting of the hyaluronate synthase gene (HAS1).
XX

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:01:16 ; Search time 3.93949 Seconds
(without alignments)
8307.054 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggcttcagagctactt 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A-COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B-COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A-COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B-COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS-COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	601	4	US-09-949-016-201780 Sequence 201780,
2	20	100.0	2087	4	US-09-949-016-5659 Sequence 5659, Ap
3	20	100.0	2116	3	US-08-635-552A-1 Sequence 1, Appli
4	20	100.0	2117	3	US-09-155-768-1 Sequence 1, Appli
5	20	100.0	14862	4	US-09-949-016-17401 Sequence 17401, A
6	17.4	87.0	91831	4	US-09-949-016-13694 Sequence 13694, A
7	16.4	82.0	3831	4	US-09-360-394C-1 Sequence 1, Appli
8	15.8	79.0	400	4	US-08-956-171E-4058 Sequence 4058, Ap
9	15.8	79.0	400	4	US-08-781-986A-4058 Sequence 4058, Ap
10	15.8	79.0	9623	4	US-08-956-171E-166 Sequence 166, App
11	15.8	79.0	9623	4	US-08-781-986A-166 Sequence 166, App
12	15.8	79.0	193169	4	US-09-949-016-15091 Sequence 15091, A
13	15.8	78.0	246240	2	US-08-724-394A-20 Sequence 20, Appl
14	15.8	79.0	246240	2	US-08-724-394A-21 Sequence 21, Appl
15	15.8	79.0	246240	2	US-08-724-394A-22 Sequence 22, Appl
16	15.4	77.0	1642	4	US-10-000-489-77 Sequence 77, Appl
17	15.4	77.0	13782	4	US-09-949-016-12802 Sequence 12802, A
18	15.2	76.0	601	4	US-09-949-016-13694 Sequence 13694, A
19	15.2	76.0	750	4	US-09-489-039A-1506 Sequence 1506, Ap
20	15.2	76.0	1844	4	US-09-949-016-5123 Sequence 5123, Ap
21	15.2	76.0	1856	1	US-08-706-214-1 Sequence 1, Appli
22	15.2	76.0	1856	4	US-09-949-016-475 Sequence 475, App
23	15.2	76.0	2492	3	US-09-327-984A-1 Sequence 1, Appli
24	15.2	76.0	17472	4	US-09-949-016-12217 Sequence 12217, A
25	15.2	76.0	17476	4	US-09-949-016-16865 Sequence 16865, A
26	15.2	76.0	19854	4	US-09-949-016-12585 Sequence 12585, A
27	15.2	76.0	19854	4	US-09-949-016-17306 Sequence 17306, A

```

28 15.2 76.0 21721 4 US-09-269-939A-41 Sequence 41, Appl
29 15.2 76.0 22976 4 US-09-269-939A-19 Sequence 19, Appl
30 15.2 76.0 23187 4 US-09-499-522-1 Sequence 1, Appli
31 15.2 76.0 97423 4 US-09-949-016-12742 Sequence 12742, A
32 15.2 76.0 97424 4 US-09-949-016-15576 Sequence 15576, A
33 15.2 76.0 108440 4 US-09-949-016-12065 Sequence 12065, A
34 15.2 76.0 108441 4 US-09-949-016-14090 Sequence 14090, A
35 15.2 76.0 1830121 4 US-09-557-884-1 Sequence 1, Appli
36 15.2 76.0 1830121 4 US-09-643-990A-1 Sequence 1, Appli
37 15 75.0 639 3 US-09-171-517B-1 Sequence 1, Appli
38 15 75.0 5351 3 US-09-221-017B-771 Sequence 771, App
39 14.8 74.0 601 4 US-09-949-016-82354 Sequence 82354, A
40 14.8 74.0 601 4 US-09-949-016-82355 Sequence 82355, A
41 14.8 74.0 601 4 US-09-949-016-82356 Sequence 82356, A
42 14.8 74.0 601 4 US-09-949-016-82357 Sequence 82357, A
43 14.8 74.0 601 4 US-09-949-016-82358 Sequence 82358, A
44 14.8 74.0 601 4 US-09-949-016-82359 Sequence 82359, A
45 14.8 74.0 601 4 US-09-949-016-85681 Sequence 85681, A

```

ALIGNMENTS

```

RESULT 1
US-09-949-016-201780/c
; Sequence 201780, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201780
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201780

Query Match 100.0% Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTCTCAGAGCTACTT 20
Db 222 CGGGCTTCTCAGAGCTACTT 203
|||||
|||||

RESULT 2
US-09-949-016-5659
; Sequence 5659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5659

Query Match      100.0%; Score 20; DB 4; Length 2087;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGCAGAGCTACTT 20
Db 912 CGGGCTTGTGCAGAGCTACTT 931

RESULT 3
US-08-635-552A-1
; Sequence 1, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
US-08-635-552A-1

Query Match      100.0%; Score 20; DB 3; Length 2116;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGCAGAGCTACTT 20
Db 912 CGGGCTTGTGCAGAGCTACTT 931

RESULT 4
US-09-155-768-1
; Sequence 1, Application US/09155768A
; Patent No. 6162508
; GENERAL INFORMATION:
; APPLICANT: SEIKAGAKU CORPORATION
; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF
; TITLE OF INVENTION: HUMAN-ORIGIN AND DNA ENCODING THE SAME
; FILE REFERENCE: TOYAM37.001APC
; CURRENT APPLICATION NUMBER: US/09/155,768A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: JP 8-084326
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: JP 8-109663
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (149)...(1777)
US-09-155-768-1

Query Match      100.0%; Score 20; DB 3; Length 2117;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGCAGAGCTACTT 20
Db 920 CGGGCTTGTGCAGAGCTACTT 939

RESULT 5
US-09-949-016-17401
; Sequence 17401, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17401
; LENGTH: 14862
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17401

Query Match      100.0%; Score 20; DB 4; Length 14862;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGCAGAGCTACTT 20
Db 8955 CGGGCTTGTGCAGAGCTACTT 8974

RESULT 6
US-09-949-016-13694
; Sequence 13694, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```


; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13694
 ; LENGTH: 91831
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-13694

Query Match 87.0%; Score 17.4; DB 4; Length 91831;
 Best Local Similarity 94.7%; Pred. No. 26;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GGGCTGTGTCAGAGCTACTT 20
 DB 55477 GGGCTGTGTCAGAGCTACTT 55495

RESULT 7
 US-09-360-394C-1
 ; Sequence 1, Application US/09360394C
 ; Patent No. 6610509
 ; GENERAL INFORMATION:
 ; APPLICANT: Sandell, Linda
 ; APPLICANT: Xie, Wei-Pen
 ; TITLE OF INVENTION: METHODS OF TARGETED EXPRESSION BY THE CD-RAP GENE PROMOTER
 ; FILE REFERENCE: BJCH 10026
 ; CURRENT APPLICATION NUMBER: US/09/360,394C
 ; CURRENT FILING DATE: 1999-07-23
 ; PRIOR APPLICATION NUMBER: 60-094,092
 ; PRIOR FILING DATE: 1998-07-24
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 3831
 ; TYPE: DNA
 ; ORGANISM: Mus musculus;
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3598)..(3598)
 ; OTHER INFORMATION: n=unknown
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3591)..(3591)
 ; OTHER INFORMATION: n=unknown
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3624)..(3624)
 ; OTHER INFORMATION: n=unknown
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3790)..(3790)
 ; OTHER INFORMATION: n=unknown
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3805)..(3805)
 ; OTHER INFORMATION: n=unknown
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3821)..(3821)
 ; OTHER INFORMATION: n=unknown
 ; US-09-360-394C-1

Query Match 82.0%; Score 16.4; DB 4; Length 3831;
 Best Local Similarity 94.4%; Pred. No. 50;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GGGCTGTGTCAGAGCTACT 19
 DB 1176 GGGCTGTGTCAGAGCTACT 1193

RESULT 8
 US-08-956-171E-4058
 ; Sequence 4058, Application US/08956171E
 ; Patent No. 6593114
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; Gil H. Choi
 ; Patrick S. Dillon
 ; Craig A. Rosen
 ; Steven C. Barash
 ; Michael R. Fannon
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5256
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/956,171E
 FILING DATE: 20-Oct-1997
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/009,861
 FILING DATE: January 5, 1996
 APPLICATION NUMBER: 08/781,986
 FILING DATE: January 3, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Mark J. Hyman
 REGISTRATION NUMBER: 46,789
 REFERENCE/DOCKET NUMBER: PB248PI
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (240) 314-1224
 TELEFAX: (301) 309-8439
 INFORMATION FOR SEQ ID NO: 4058:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 400 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 4058:
 US-08-956-171E-4058

Query Match 79.0%; Score 15.8; DB 4; Length 400;
 Best Local Similarity 89.5%; Pred. No. 68;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GGGCTGTGTCAGAGCTACT 19
 DB 66 GGGCTGTGTCAGAGCTACT 84

RESULT 9
 US-08-781-986A-4058
 ; Sequence 4058, Application US/08781986A
 ; Patent No. 6737248
 ; GENERAL INFORMATION:

```
;
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 4058:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-781-986A-4058
;
; Query Match 79.0%; Score 15.8; DB 4; Length 400;
; Best Local Similarity 89.5%; Pred. No. 68;
; Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; Qy 1 CGGGCTTGTCTCAGCGTACT 19
; Db 66 CGGGCTTGTCTCAGCGTACT 84
;
; RESULT 10
; US-08-956-171E-166
; Sequence 166, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
;
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
```

```
;
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 166:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9623 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 166:
;
; US-08-956-171E-166
;
; Query Match 79.0%; Score 15.8; DB 4; Length 9623;
; Best Local Similarity 89.5%; Pred. No. 1.3e+02;
; Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; Qy 1 CGGGCTTGTCTCAGCGTACT 19
; Db 7504 CGGGCTTGTCTCAGCGTACT 7522
;
; RESULT 11
; US-08-781-986A-166
; Sequence 166, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 166:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9623 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-781-986A-166
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
;
US-08-724-394A-20

Query Match 79.0%; Score 15.8; DB 2; Length 246240;
Best Local Similarity 89.5%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0

QY 2 GGGCTTGTCTCAGAGCTACTT 20
|||
Db 60431 GGGCTTGTCTCAGAGGTATT 60449

RESULT 14
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Laufer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereo
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300

```

; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match 79.0%; Score 15.8; DB 2; Length 246240;
Best Local Similarity 89.5%; Pred.No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCTTGTCTCAGAGCTACTT 20
|||||
Db 60431 GGGCTTGTCTCAGAGGTATT 60449

RESULT 15
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22

Query Match 79.0%; Score 15.8; DB 2; Length 246240;
Best Local Similarity 89.5%; Pred.No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GGGCTTGTCTCAGAGCTACTT 20
|||||
Db 60431 GGGCTTGTCTCAGAGGTATT 60449
Search completed: March 13, 2005, 08:12:26
Job time : 7.93949 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:08:52 ; Search time 82.4452 Seconds
(without alignments)
9233.835 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggcttcagagctactt 20

Scoring table: IDENTITY_NUC

Gapop 10_0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_ges1:
9: gb_ges2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	551	4	BG707429
2	20	100.0	551	4	BG707608
3	20	100.0	610	5	BX327795
4	20	100.0	739	9	AY421570
5	20	100.0	782	4	BI753116
6	20	100.0	896	5	BX371636
7	20	100.0	1084	4	BM544718
8	20	100.0	1728	9	AY421569
9	20	100.0	2072	3	CR602106
10	18.4	92.0	166	5	BM936724
11	18.4	92.0	495	7	CN664556
12	18.4	92.0	1746	9	AY421571
13	18.4	92.0	3542	3	AK053726
14	17.4	87.0	442	5	BX264217
15	17.4	87.0	500	7	CK675190
16	17.4	87.0	502	5	BX372063
17	17.4	87.0	506	5	BQ169305
18	17.4	87.0	514	1	AL927904
19	17.4	87.0	570	1	A1878069
20	17.4	87.0	644	7	CF253438
21	17.4	87.0	681	6	CD777815
22	17.4	87.0	686	5	BX301603
23	17.4	87.0	722	5	BX306034
24	17.4	87.0	723	5	BX302845

C	25	17.4	87.0	743	5	BU269373	BU269373	603816334
	26	17.4	87.0	754	6	CD777811	CD777811	TDSUBS RF
	27	17.4	87.0	759	7	CK687847	CK687847	ZF101-P00
C	28	17.4	87.0	855	1	AJ444643	AJ444643	AG444643
	29	17.4	87.0	868	7	CO924080	AG68COURT	AG68COURT
	30	17.4	87.0	888	7	CK017913	CK017913	AGENCOURT
C	31	17.4	87.0	891	5	BU230362	BU230362	603400985
	32	17.4	87.0	1018	5	BQ891951	BQ891951	AGENCOURT
C	33	17	85.0	336	7	CK116412	CK116412	B011P46 H
C	34	17	85.0	456	7	CK116743	CK116743	B016P58 H
C	35	17	85.0	514	1	AI162566	AI162566	A019P49U
C	36	17	85.0	721	5	BU884547	BU884547	R012B07 P
	37	16.8	84.0	356	2	BF708639	BF708639	MI-P-AY0-
C	38	16.8	84.0	437	9	FR002608	286391 F.rubripes	BA008
C	39	16.8	84.0	619	9	CL588198	CL588198	OB Ba008
C	40	16.8	84.0	622	2	BE289996	BE289996	60T089569
C	41	16.8	84.0	736	9	AG411055	AG411055	MUG mubcu
C	42	16.8	84.0	770	9	CL516805	CL516805	SAIL_95_D
	43	16.8	84.0	817	5	BU451953	BU451953	603215232
C	44	16.8	84.0	1036	9	CNS04RFO	AA303813	Tetracdon
C	45	16.4	82.0	303	1	AA781062	AA781062	a110g09.s

ALIGNMENTS

RESULT 1
BG707429
LOCUS
DEFINITION
602672848F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795643 5', mRNA linear EST 07-MAY-2001
mRNA sequence.
BG707429
ACCESSION
BG707429
VERSION
BG707429.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 551)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0678 row: p column: 12
High quality sequence stop: 551.
Location/Qualifiers
1. 551

FEATURES

source
1. 551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4795643"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH MGC 96"
/note="Organ: brain; Vector: pbluescriptR (modified pbluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

```

ORIGIN
Query Match      100.0%; Score 20; DB 4; Length 551;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||
Db 248 CGGGCTTGTCTCAGAGCTACTT 267

RESULT 2
LOCUS BG707608 551 bp mRNA linear EST 07-MAY-2001
DEFINITION 602670672F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793704 5',
            mRNA sequence.
ACCESSION BG707608
VERSION BG707608.1 GI:13984125
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 551)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30340686.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0BAG028ZH06_CS03640_1&c=9502.r

FEATURES
            source
            Location/Qualifiers
                1..610
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DI027YJ02"
                /tissue_type="PLACENTA COT 25-NORMALIZED"
                /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                /note="1st strand cDNA was primed with a NotI-oligo(dT)
                primer. Five prime end enriched, double-strand cDNA was
                digested with Not I and cloned into the Not I and EcoR V
                sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      100.0%; Score 20; DB 5; Length 610;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||
Db 219 CGGGCTTGTCTCAGAGCTACTT 238

RESULT 4
LOCUS AY421570 739 bp DNA linear GSS 17-DEC-2003
DEFINITION Pan troglodytes HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION AY421570
VERSION AY421570.1 GI:39748429
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 739)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (552), 1960-1963 (2003)
14671302
2 (bases 1 to 739)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,

ORIGIN
Query Match      100.0%; Score 20; DB 4; Length 551;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||
Db 248 CGGGCTTGTCTCAGAGCTACTT 267

RESULT 3
LOCUS BX327795 610 bp mRNA linear EST 08-APR-2004
DEFINITION BX327795 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

```

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sniinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES

source

1..739
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>739
/gene="HAS1"
/locus_tag="HCM7601"

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 739;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTACAGCTACTT 20
|||||
Db 132 CGGGCTTGTACAGCTACTT 151

RESULT 5

BI753116

LOCUS

DEFINITION 603025961F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196399 5',
mRNA sequence.

ACCESSION

BI753116

VERSION

BI753116.1

KEYWORDS

EST.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 782)

NIH-MGC

http://mgs.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM1491 row: j column: 16

High quality sequence stop: 780.

Location/Qualifiers

1..782

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5196399"

/lab_host="DH10B"

/clone_lib="NIH_MGC_114"

/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: EcoRV (destroyed); RNA source anonymous pool of 6

male brains, age range 23-27 yo. Library is oligo-dT

primed and directionally cloned (EcoRV site is destroyed

upon cloning). Average insert size 1.5 kb, insert size

range 1-3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 019. Note:

this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 20; DB 4; Length 782;

Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTACAGCTACTT 20
|||||

Db 212 CGGGCTTGTACAGCTACTT 231
|||||

RESULT 6

BX371636/c

LOCUS

DEFINITION BX371636 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI027YJ02 3-PRIME, mRNA sequence.

ACCESSION

BX371636

VERSION

BX371636.2

KEYWORDS

EST.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 896)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 8, 2003 this sequence version replaced gi:30450029.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

9502.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?g=CS0BAI019ZC12_CS01819_1&c=9502.r

FEATURES

source

1..896
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI027YJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 896;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTACAGCTACTT 20
|||||

Db 356 CGGGCTTGTACAGCTACTT 337
|||||

RESULT 7

BM544718

LOCUS

DEFINITION BM544718 1084 bp mRNA linear EST 20-FEB-2002
AGENCOURT_6494603 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5727446
5', mRNA sequence.

ACCESSION

BM544718

VERSION

BM544718.1

KEYWORDS

EST.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM12721 row: i column: 15
High quality sequence start: 5
High quality sequence stop: 689.
Location/Qualifiers
1. .1084
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5727446"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH MGC 124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
(destroyed); Site_2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned [EcoRV site is destroyed upon cloning]. Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."

ORIGIN
Query Match 100.0%; Score 20; DB 4; Length 1084;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGTACTT 20
|||||
Db 276 CGGGCTTGTCTCAGAGTACTT 295

RESULT 8
AY421569
LOCUS AY421569 1728 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY421569
VERSION AY421569.1 GI:39748428
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1728)
Clark.A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1728)
Clark.A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1. .1728
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1728
/genes="HAS1"
/locus_tag="HGM7601"

ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 1728;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGTACTT 20
|||||
Db 868 CGGGCTTGTCTCAGAGTACTT 887

RESULT 9
CR602106
LOCUS CR602106 2072 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DI027YJ02 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR602106
VERSION CR602106.1 GI:50482913
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2072)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 2072)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .2072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI027YJ02"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 100.0%; Score 20; DB 3; Length 2072;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGTACTT 20
|||||
Db 925 CGGGCTTGTCTCAGAGTACTT 944

RESULT 10
BM936724
LOCUS BM936724 166 bp mRNA linear EST 13-MAR-2002

```


DEFINITION UI-M-BH3-ar-d-04-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
 UI-M-BH3-ar-d-04-0-UI.5', mRNA sequence.
 ACCESSION BM936724
 VERSION BM936724.1 GI:19395876
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 166)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 9704477
 PUBMED 8889548
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 REVERSE.

FEATURES

Location/Qualifiers
 1..166
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-ar-d-04-0-UI"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH_BMAP_M_S4"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; The
 NIH_BMAP_M_S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH_BMAP_M_S4,
 NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
 NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
 (NIH_BMAP_M_S4) was constructed as follows: PCR amplified
 cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
 NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH_BMAP_M_S4 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)."

ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 166;
 Best Local Similarity 95.0%; Pred. No. 66;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGAGCTTGTACAGCTACTT 20
 |||||||||||||||||||

Db 91 CGAGCTTGTACAGCTACTT 110

RESULT 11
 CN664556/c

LOCUS
 DEFINITION

ACCESSION
 VERSION

KEYWORDS
 SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CN664556 495 bp mRNA linear EST 17-MAY-2004
 A0814F02-5 NIA Mouse E13.5 whole embryo cDNA library (long) Mus
 musculus cDNA clone NIA:A0814F02 IMAGE:30758557 5', mRNA sequence.
 CN664556
 CN664556.1 GI:47431007
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 495)
 Sharov,A.A., Piao,Y., Matoba,R., Dudekula,D.B., Qian,Y.,
 VanBuren,V., Falco,G., Martin,P.R., Stagg,C.A., Bassey,U.C.,
 Wang,Y., Carter,M.G., Hamatani,T., Aiba,K., Akutsu,H., Sharova,L.,
 Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S.,
 Nagaraja,R., Boheler,K.R., Taub,D., Hodes,R.J., Longo,D.L.,
 Schlessinger,D., Keller,J., Klotz,E., Kelsoe,G., Umezawa,A.,
 Vescovi,A.L., Rosant,J., Kunath,T., Hogan,B.L., Curci,A.,
 D'Urso,M., Kelsoe,J., Hide,W. and Ko,M.S.
 Transcriptome analysis of mouse stem cells and early embryos
 PLoS Biol. 1 (3), 410-419 (2003)
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 Plate: A0814 row: F column: 02
 Seq primer: M13 Reverse
 High quality sequence stop: 495
 POLYA=No.

FEATURES
 source

Location/Qualifiers
 1..495
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="niaEST:A0814F02-5"
 /db_xref="taxon:10090"
 /clone="NIA:A0814F02 IMAGE:30758557"
 /tissue="type=whole embryo including extraembryonic
 tissues at 13.5-days postcoitum"
 /dev_stage="E13.5"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse E13.5 whole embryo cDNA library
 (long)"
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
 Site 2: NotI; Mouse cDNA project by the Laboratory of
 Genetics, National Institute on Aging (NIA), Intramural
 Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA).
 This is a long-transcript enriched cDNA library (Ref.
 Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
 RNAs were extracted from 1 embryo at 13.5-days postcoitum.
 Double-stranded cDNAs were synthesized with an Oligo(dT)
 primer [Invitrogen:
 5'-PGACTAGTCTAGATCGCGAGCGCCCTTTT-3'] from
 3ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Lona-linker LK-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
 vector. The DH10B E. coli host was transformed with the
 ligation mixture by the standard chemical method. The
 average insert size is about 3.0Kb. The library was
 constructed by Yulan Piao."

ORIGIN

Query Match 92.0%; Score 18.4; DB 7; Length 495;
 Best Local Similarity 95.0%; Pred. No. 81;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGCTTGTGCAGAGCTACTT 20

DB 65 CGAGCTTGTGCAGAGCTACTT 46

RESULT 12

AY421571

LOCUS

DEFINITION Mus musculus HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence, GSS 17-DEC-2003

ACCESSION AY421571

VERSION AY421571.1 GI:39748430

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE

AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

Source

1..1746

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

<1..>1746

/gene="HAS1"

/locus_tag="HCM7601"

ORIGIN

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 9; Length 1746;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

1 CGGGCTTGTGCAGAGCTACTT 20

DB 886 CGAGCTTGTGCAGAGCTACTT 905

RESULT 13

AK053726

LOCUS

DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:E130302P16 product:hyaluronan synthase1, full insert sequence.

ACCESSION AK053726

VERSION AK053726.1 GI:26343670

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

6 (bases 1 to 3542)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

Location/Qualifiers

1..3542

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:E130302P16"

/db_xref="taxon:10090"
/clone="E130302P16"
/tissue_type="eyeball"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
61..456

CDS

/note="unnamed protein product; hyaluronan synthase1
(MGDI|GI:106590, GB|NM_008215, evidence: BLASTN, 99%,
match=762)
/codon_start=1
/protein_id="BAC35492.1"
/db_xref="GI:26343671"
/translation="NRQDMPKPSAARCCSGIARRALTIIFALLILGLMTWAYAGVP
LASDRYGLAFGLYGAFLSAHLVAOSLFAYLEHRRVAAARRSLAKGPLDAATARSVA
LTISAYQBDPAVLRLQCLISARALLYPHIR"

ORIGIN

Query Match 92.0%; Score 18.4; DB 3; Length 3542;
Best Local Similarity 95.0%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGCTTGTTCAGAGCTACTT 20
DB 2519 CGAGCTTGTTCAGAGCTACTT 2538

RESULT 14
BX264217/c
LOCUS
DEFINITION
BX264217 AGENAE Gallus gallus multi-tissues normalized and
once-subtracted cDNA library (gcal) Gallus gallus cDNA clone
gcal0014c.h.13 5prim, mRNA sequence.
ACCESSION
BX264217
VERSION
BX264217.1 GI:28586815
KEYWORDS
EST.
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Herauld, F., Le Muth-Metzinger, V., Desert, C., Retout, E., Piumi, F.,
Klopp, C. and Douaire, M.
1 (bases 1 to 442)
Construction and primary characterization of chicken normalized
multi-tissue cDNA libraries
Unpublished (2003)
Contact: Douaire M
INRA, UMR INRA-ENSAR Genetique Animale
65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE
Tel: +33 (0) 2.23.48.54.63
Fax: +33 (0) 2.23.48.54.70

Email: Madeleine.Douaire@roazhon.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at signasupport@jey.inra.fr to obtain the chromatogram of this
sequence.
Plate: 0014 row: h column: 13
Seq primer: M13R.
Location/Qualifiers
1..442
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="gcal0014c.h.13"
/tissue_type="multi-tissues"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="AGENAE Gallus gallus multi-tissues normalized
and once-subtracted cDNA library (gcal)"
/note="vector: pTV3-pac; tissues: adipose tissue, brain,
kidney, liver, multi-tissues, muscle, ovary, testis, bone
marrow, caecum, duodenum, embryos, jejunum, ileon, oviduct,
granulosa, hypothalamus, ileon, jejunum, oviduct,

FEATURES
source

1..442
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="gcal0014c.h.13"
/tissue_type="multi-tissues"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="AGENAE Gallus gallus multi-tissues normalized
and once-subtracted cDNA library (gcal)"
/note="vector: pTV3-pac; tissues: adipose tissue, brain,
kidney, liver, multi-tissues, muscle, ovary, testis, bone
marrow, caecum, duodenum, embryos, jejunum, ileon, oviduct,

pancreas, skin, spleen, thymus, utero-vaginal gland,
pituitary gland, hematopoietic progenitor cells, small
follicle. Clone distribution : AGENAE Resource centre.
Francois Piumi, Francois.Piumi@inra.fr, INRA, CEA
Radiobiologie et Etude du genome (LREG), Domaine de
Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"

ORIGIN

Query Match 87.0%; Score 17.4; DB 5; Length 442;
Best Local Similarity 94.7%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGCTTGTTCAGAGCTACTT 19
DB 216 CGGCTTGTTCAGAGCTACTT 198

RESULT 15

CK675190
LOCUS
DEFINITION
ZF101-P00004-DPE-F12 GISZF001-ra Danio rerio cDNA clone
IMAGE:6912493 5', mRNA sequence.

ACCESSION
CK675190
VERSION
CK675190.1 GI:42427342
KEYWORDS
EST.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 500)
AUTHORS
Wei, C., Mathavan, S., Thoreau, H., Lim, L., Lee, C. and Ruan, Y.
Genome Institute of Singapore, Zebrafish Gene Collection
Unpublished (2004)
Contact: Ruan Y

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Cloning and Sequencing
Genome Institute of Singapore
60 Biopolis Street, #02-01, Genome, Singapore 138672
Tel: +65 6478 8073
Fax: +65 6478 9059
Email: ruanyj@gis.a-star.edu.sg
GIS Clone ID: ZF101-P00004-BR2_F12
PCR Primers
FORWARD: M13
BACKWARD: M13
Plate: ZF101-P00004-BR2 row: F column: 12
Seq primer: CCGCATAACTGTATAGCA
High quality sequence stop: 500.
Location/Qualifiers
1..500
/organism="Danio rerio"
/mol_type="mRNA"
/strain="Singapore local strain"
/db_xref="taxon:7955"
/clone="IMAGE:6912493"
/tissue_type="Embryo"
/dev_stage="7 Different embryonic stages (From just
fertilized Embryos to 72 hours just hatched baby fish)"
/lab_host="DH10B"
/clone_lib="GISZF001_ra"
/note="Vector: pDNR-LiB; Site 1: Sfi A (GCCATTACGGCC);
Site 2: Sfi B (GGCGAGGGCGCC); Priming method: Sfi-(dT)30
Primed; Priming sequence:
5.ATTCTAGAGCGGAGCGGCGGCGGACATG(T)30VN ; Directionally
cloned, 5' cloning site: Sfi A site GCCATTACGGCC ; 5'
linker/adaptor sequence: 5.AAGCAGTGGTATCAACGAGAGTGGCC ;
3' cloning site: Sfi B site GGCGAGGGCGCC ; 3'
linker/adaptor sequence: same as the priming sequence ;
Average insert size: 2kb ; For PCR insert analysis: Use
M13 Forward and reverse primers ; Library Amplified ;
Recombinants (inserts): 98% ; Library complexity: 5x10⁶ ;
Full-length construction (method): SMART, a Clontech
method The pooled tissue RNA was collected and used to

FEATURES
source

1..500
/organism="Danio rerio"
/mol_type="mRNA"
/strain="Singapore local strain"
/db_xref="taxon:7955"
/clone="IMAGE:6912493"
/tissue_type="Embryo"
/dev_stage="7 Different embryonic stages (From just
fertilized Embryos to 72 hours just hatched baby fish)"
/lab_host="DH10B"
/clone_lib="GISZF001_ra"
/note="Vector: pDNR-LiB; Site 1: Sfi A (GCCATTACGGCC);
Site 2: Sfi B (GGCGAGGGCGCC); Priming method: Sfi-(dT)30
Primed; Priming sequence:
5.ATTCTAGAGCGGAGCGGCGGCGGACATG(T)30VN ; Directionally
cloned, 5' cloning site: Sfi A site GCCATTACGGCC ; 5'
linker/adaptor sequence: 5.AAGCAGTGGTATCAACGAGAGTGGCC ;
3' cloning site: Sfi B site GGCGAGGGCGCC ; 3'
linker/adaptor sequence: same as the priming sequence ;
Average insert size: 2kb ; For PCR insert analysis: Use
M13 Forward and reverse primers ; Library Amplified ;
Recombinants (inserts): 98% ; Library complexity: 5x10⁶ ;
Full-length construction (method): SMART, a Clontech
method The pooled tissue RNA was collected and used to

construct full length enriched cDNA library and also served as template to synthesize complex first strand cDNA probe. Two high density colony arrays were made from over 110K cDNA clones and hybridized with the probes. Low intensity clones were selected as they represented rare expressed clones. The hybridization intensities for all clones span from 0 to 1.8 million counts and the low abundant class ranged from 0 to 13,000."

ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 500;
 Best Local Similarity 94.7%; Pred. No. 2.9e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCGGCTTGTTCAGAGCTACT 19
 |||||
 Db 480 CCGGCTTGTTCAGAGCTTCT 498

Search completed: March 13, 2005, 11:55:05
 Job time : 87.4452 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 08:01:42 ; Search time 14.0784 Seconds
(without alignments)
8450.560 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggcttgctcagactactt 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	18	US-10-672-399-9
2	20	100.0	490	10	US-09-918-995-27210
3	20	100.0	1065	18	US-10-672-399-7
4	20	100.0	1071	18	US-10-672-399-3
5	20	100.0	1083	18	US-10-672-399-5
6	20	100.0	1073	18	US-10-672-399-1
7	20	100.0	2116	13	US-10-042-523-1
8	18.4	92.0	1752	10	US-09-902-939-1
9	16.4	82.0	575	18	US-10-425-115-53796
10	16.4	82.0	2928	18	US-10-437-963-33800
11	16.4	82.0	21410	13	US-10-087-192-1231

12	16.4	82.0	30486	17	US-10-085-117-19
13	15.8	79.0	201	19	US-10-741-600-37582
14	15.8	79.0	213	18	US-10-425-115-167738
15	15.8	79.0	316	11	US-09-732-627A-2660
16	15.8	79.0	400	8	US-08-781-986A-4058
17	15.8	79.0	400	17	US-10-329-624-4058
18	15.8	79.0	477	9	US-09-864-761-1728
19	15.8	79.0	582	9	US-09-864-761-12208
20	15.8	79.0	913	13	US-10-027-632-161814
21	15.8	79.0	913	17	US-10-027-632-161814
22	15.8	79.0	1607	17	US-10-425-114-13595
23	15.8	79.0	1953	17	US-10-282-122A-7912
24	15.8	79.0	3017	18	US-10-425-115-104761
25	15.8	79.0	9623	8	US-08-781-986A-166
26	15.8	79.0	9623	17	US-10-329-624-166
27	15.8	79.0	43975	19	US-10-741-600-17669
28	15.8	79.0	50720	18	US-10-322-281-621
29	15.8	79.0	77834	13	US-10-087-192-343
30	15.8	79.0	115863	18	US-10-723-860-2504
31	15.8	79.0	235033	15	US-10-301-844-1
32	15.8	79.0	237326	15	US-10-301-844-2
33	15.4	77.0	598	13	US-10-027-632-213175
34	15.4	77.0	598	17	US-10-027-632-213175
35	15.4	77.0	1211	17	US-10-260-238-308
36	15.4	77.0	1642	10	US-09-992-600A-77
37	15.4	77.0	1642	10	US-09-924-340-77
38	15.4	77.0	1642	10	US-09-992-095B-77
39	15.4	77.0	1642	10	US-09-999-570-77
40	15.4	77.0	1642	14	US-10-000-489-77
41	15.4	77.0	1642	14	US-10-000-986-77
42	15.4	77.0	1642	16	US-10-154-678-77
43	15.4	77.0	1642	16	US-10-001-142-77
44	15.4	77.0	1642	19	US-10-838-854-77
45	15.4	77.0	1655	18	US-10-437-963-97018

ALIGNMENTS

RESULT 1
US-10-672-399-9
; Sequence 9, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-9

Query Match 100.0%; Score 20; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGGCTTGTCTCAGAGCTACTT 20
Db 1 CCGGCTTGTCTCAGAGCTACTT 20

RESULT 2
US-09-918-995-27210
; Sequence 27210, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:

```
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27210

Query Match      100.0%; Score 20; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 335 CGGGCTTGTCTCAGAGCTACTT 354

RESULT 3
US-10-672-399-7
; Sequence 7, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-7

Query Match      100.0%; Score 20; DB 18; Length 1065;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 877 CGGGCTTGTCTCAGAGCTACTT 896

RESULT 4
US-10-672-399-3
; Sequence 3, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3

Query Match      100.0%; Score 20; DB 18; Length 1065;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 877 CGGGCTTGTCTCAGAGCTACTT 896

RESULT 5
US-10-672-399-5
; Sequence 5, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-5

Query Match      100.0%; Score 20; DB 18; Length 1083;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 877 CGGGCTTGTCTCAGAGCTACTT 896

RESULT 6
US-10-672-399-1
; Sequence 1, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-1

Query Match      100.0%; Score 20; DB 18; Length 1737;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 877 CGGGCTTGTCTCAGAGCTACTT 896

RESULT 7
US-10-672-399-1
; Sequence 1, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-1

Query Match      100.0%; Score 20; DB 18; Length 1737;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 877 CGGGCTTGTCTCAGAGCTACTT 896
```

US-10-042-523-1
; Sequence 1, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briekin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-042-523-1
Query Match 100.0%; Score 20; DB 13; Length 2116;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 912 CGGGCTTGTCTCAGAGCTACTT 931
RESULT 8
; Sequence 1, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHaza
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-US0
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: mouse

US-09-902-939-1
Query Match 92.0%; Score 18.4; DB 10; Length 1752;
Best Local Similarity 95.0%; Pred. No. 1.1;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 892 CGAGCTTGTCTCAGAGCTACTT 911
RESULT 9
US-10-425-115-53796
; Sequence 53796, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 53796
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_149063C.1
US-10-425-115-53796
Query Match 82.0%; Score 16.4; DB 18; Length 575;
Best Local Similarity 94.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 GGCTTGTCTCAGAGCTACTT 20
Db 480 GGCTTGTCTCAGAGCTACTT 497
RESULT 10
US-10-437-963-33800/c
; Sequence 33800, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 33800
; LENGTH: 2928
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37878C.1
US-10-437-963-33800
Query Match 82.0%; Score 16.4; DB 18; Length 2928;
Best Local Similarity 94.4%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY      2 GGGCTTGTCTCAGAGCTACT 19
      |||||
Db      2706 GGGCTTGTCTCAGAGCTACT 2689
      |||||

RESULT 11
US-10-087-192-1231
; Sequence 1231, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1231
; LENGTH: 21410
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(21410)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1231

Query Match      82.0%; Score 16.4; DB 13; Length 21410;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GGGCTTGTCTCAGAGCTACT 19
      |||||
Db      9034 GGGCTTGTCTCAGAGCTACT 9051
      |||||

RESULT 12
US-10-085-117-19
; Sequence 19, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 30486
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(30486)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-19

Query Match      82.0%; Score 16.4; DB 17; Length 30486;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GGGCTTGTCTCAGAGCTACT 19
      |||||
Db      9034 GGGCTTGTCTCAGAGCTACT 9051
      |||||

RESULT 13
US-10-741-600-37582/c
; Sequence 37582, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37582
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-37582

Query Match      79.0%; Score 15.8; DB 19; Length 201;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CGGCTTGTCTCAGAGCTACT 19
      |||||
Db      24 CGGCTTGTCTCAGAGCTACT 6
      |||||

RESULT 14
US-10-425-115-167738
; Sequence 167738, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 167738
; LENGTH: 213
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_84558C.1
US-10-425-115-167738

Query Match      79.0%; Score 15.8; DB 18; Length 213;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GGGCTTGTCTCAGAGCTACTT 20
      |||||
Db      123 GGGCTTGTCTCAGAGCTACTT 141
      |||||

RESULT 15
US-09-732-627A-2660
; Sequence 2660, Application US/09732627A
; Publication No. US20040123338A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen L.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(51770)B
US-09-732-627A-2660
```



```

; CURRENT APPLICATION NUMBER: US/09/732,627A
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4930
; SEQ ID NO 2660
; LENGTH: 316
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(316)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3493-007-P1-M1-C5
US-09-732-627A-2660

```

```

Query Match      79.0%; Score 15.8; DB 11; Length 316;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 CGGGCTTCTCAGAGCTACT 19
         |||||
Db      56 CGGGCTTCTCAGAGCTGCT 74

```

Search completed: March 13, 2005, 17:10:04
Job time : 15.0784 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:11:12 ; Search time 97.2028 Seconds
(without alignments)
8399.558 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggctgtgcagactact 20

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_Main:

- 1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
- 12: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
- 13: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
- 14: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
- 15: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
- 16: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
- 17: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
- 18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq.*
- 19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq.*
- 20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq.*
- 22: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
- 23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
- 24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
- 25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
- 26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
- 27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
- 28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
- 29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
- 30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
- 31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
- 32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
- 33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
- 34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
- 35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
- 36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
- 37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
- 38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq.*
- 39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
- 40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
- 41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
- 42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
- 43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*

- 44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*
- 45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq.*
- 46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
- 47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
- 48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
- 49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
- 50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
- 51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
- 52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
- 53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
- 54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
- 55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
- 56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq.*
- 57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq.*
- 58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq.*
- 59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq.*
- 60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq.*
- 61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq.*
- 62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq.*
- 63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq.*
- 64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq.*
- 65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq.*
- 66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq.*
- 67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq.*
- 68: /cgn2_6/ptodata/1/pna/US110_COMB.seq.*
- 69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
- 70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
- 71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
- 72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
- 73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
- 74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
- 75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
- 76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
- 77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
- 78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
- 79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
- 80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
- 81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
- 82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
- 83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
- 84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
- 85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
- 86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
- 87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
- 88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
- 89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
- 90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
- 91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
- 92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq.*
- 93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
- 94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
- 95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
- 96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
- 97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
- 98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
- 99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
- 100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
- 101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
- 102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
- 103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
- 104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
- 105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
- 106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
- 107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
- 108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
- 109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
- 110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
- 111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
- 112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
- 113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
- 114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
- 115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
- 116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*

117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*
 118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq.*
 119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq.*
 120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq.*
 121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq.*
 122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq.*
 123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq.*
 124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq.*
 125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq.*
 126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq.*
 127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq.*
 128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq.*
 129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	58	US-10-672-399-9	Sequence 9, Appli
2	20	100.0	201	67	US-10-990-328-92497	Sequence 92497, A
3	20	100.0	201	67	US-10-990-328-92508	Sequence 92508, A
4	20	100.0	201	67	US-10-990-328-410752	Sequence 410752, A
5	20	100.0	201	115	US-60-452-680-115499	Sequence 115499, A
6	20	100.0	201	115	US-60-453-050-82424	Sequence 82424, A
7	20	100.0	201	115	US-60-453-135-82424	Sequence 82424, A
8	20	100.0	201	116	US-60-466-412-82424	Sequence 82424, A
9	20	100.0	201	116	US-60-466-412-402528	Sequence 402528, A
10	20	100.0	490	18	US-09-235-076-27210	Sequence 27210, A
11	20	100.0	490	19	US-09-289-768-29320	Sequence 29320, A
12	20	100.0	490	20	US-09-332-782-27210	Sequence 27210, A
13	20	100.0	490	33	US-09-737-223-27210	Sequence 27210, A
14	20	100.0	490	39	US-09-918-995-27210	Sequence 27210, A
15	20	100.0	490	40	US-09-939-397-29320	Sequence 29320, A
16	20	100.0	601	66	US-10-940-774-201780	Sequence 201780, A
17	20	100.0	1065	58	US-10-672-399-7	Sequence 7, Appli
18	20	100.0	1071	58	US-10-672-399-3	Sequence 3, Appli
19	20	100.0	1083	58	US-10-672-399-5	Sequence 5, Appli
20	20	100.0	1737	1	PCT-US02-41225A-223	Sequence 223, App
21	20	100.0	1737	1	PCT-US02-41225A-224	Sequence 224, App
22	20	100.0	1737	58	US-10-672-399-1	Sequence 1, Appli
23	20	100.0	1737	95	US-60-257-537-223	Sequence 223, App
24	20	100.0	1737	95	US-60-257-537-224	Sequence 224, App
25	20	100.0	2087	1	PCT-US02-41225A-226	Sequence 226, App
26	20	100.0	2087	49	US-10-170-235-17298	Sequence 37298, A
27	20	100.0	2087	66	US-10-940-774-5659	Sequence 5659, App
28	20	100.0	2087	95	US-60-257-537-226	Sequence 226, App
29	20	100.0	2087	115	US-60-452-680-12253	Sequence 12253, A
30	20	100.0	2087	115	US-60-453-050-7646	Sequence 7646, Ap
31	20	100.0	2087	115	US-60-453-135-7646	Sequence 7646, Ap
32	20	100.0	2087	116	US-60-466-412-7646	Sequence 7646, Ap
33	20	100.0	2088	1	PCT-US02-41225A-222	Sequence 222, App
34	20	100.0	2088	52	US-10-342-164-1	Sequence 1, Appli
35	20	100.0	2088	95	US-60-257-537-222	Sequence 222, App
36	20	100.0	2089	67	US-10-990-328-7459	Sequence 7459, Ap
37	20	100.0	2098	97	US-60-278-258-16271	Sequence 16271, A
38	20	100.0	2098	97	US-10-990-328-7460	Sequence 7460, Ap
39	20	100.0	2112	102	US-60-324-185-23093	Sequence 23093, A
40	20	100.0	2116	46	US-10-042-523-1	Sequence 1, Appli
41	20	100.0	14862	66	US-10-940-774-17401	Sequence 17401, A
42	20	100.0	22862	116	US-60-466-412-87508	Sequence 87508, A
43	20	100.0	22877	67	US-10-990-328-96152	Sequence 96152, A
44	20	100.0	231222	1	PCT-US02-41225A-225	Sequence 225, App
45	20	100.0	231222	95	US-60-257-537-225	Sequence 225, App

ALIGNMENTS

```

RESULT 1
US-10-672-399-9
; Sequence 9, Application US/10672399
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672.399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-9

Query Match      100.0%; Score 20; DB 58; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CGGGCTTGTGACAGCTACTT 20
      |||
Db      1 CGGGCTTGTGACAGCTACTT 20

RESULT 2
US-10-990-328-92497
; Sequence 92497, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990.328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92497
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-990-328-92497

Query Match      100.0%; Score 20; DB 67; Length 201;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CGGGCTTGTGACAGCTACTT 20
      |||
Db     180 CGGGCTTGTGACAGCTACTT 199

RESULT 3
US-10-990-328-92508
; Sequence 92508, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990.328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92508
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens

```

RESULT 6
US-60-453-050-82424
; Sequence 82424, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May

```
Query Match      100.0%; Score 20; DB 116; Length 201;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||||||||||||||||
Db 180 CGGGCTTGTCTCAGAGCTACTT 199

RESULT 9
US-60-466-412-402528
; Sequence 402528, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: IAKOUBOVA, Michela
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01466
; CURRENT APPLICATION NUMBER: US/60/466.412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 402528
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-466-412-402528

Query Match 100.0%; Score 20; DB 116; Length 201;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||||||||||||||||
Db 180 CGGGCTTGTCTCAGAGCTACTT 199

RESULT 10
US-09-235-076-27210
; Sequence 27210, Application US/09235076
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/235,076
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-235-076-27210

Query Match 100.0%; Score 20; DB 18; Length 490;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||||||||||||||||
Db 335 CGGGCTTGTCTCAGAGCTACTT 354

RESULT 11
US-09-289-768-29320
; Sequence 29320, Application US/09289768
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-765
; CURRENT APPLICATION NUMBER: US/09/289,768
```

```
; CURRENT FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 39996
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29320
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-289-768-29320

Query Match 100.0%; Score 20; DB 19; Length 490;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||||||||||||||||
Db 335 CGGGCTTGTCTCAGAGCTACTT 354

RESULT 12
US-09-332-782-27210
; Sequence 27210, Application US/09332782A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/332,782A
; CURRENT FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-332-782-27210

Query Match 100.0%; Score 20; DB 20; Length 490;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||||||||||||||||
Db 335 CGGGCTTGTCTCAGAGCTACTT 354

RESULT 13
US-09-737-223-27210
; Sequence 27210, Application US/09737223
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/737,223
; CURRENT FILING DATE: 2000-12-13
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-737-223-27210

Query Match 100.0%; Score 20; DB 33; Length 490;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGGGCTTGTACAGCTACTT 20
|||||
Db 335 CGGGCTTGTACAGCTACTT 354

RESULT 14

US-09-918-995-27210
; Sequence 27210, Application US/09918995
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27210

Query Match 100.0%; Score 20; DB 39; Length 490;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGGGCTTGTACAGCTACTT 20
|||||
Db 335 CGGGCTTGTACAGCTACTT 354

RESULT 15

US-09-939-397-29320
; Sequence 29320, Application US/09939397
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-765
; CURRENT APPLICATION NUMBER: US/09/939,397
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 09/289,768
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 39996
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29320
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-939-397-29320

Query Match 100.0%; Score 20; DB 40; Length 490;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGGGCTTGTACAGCTACTT 20
|||||
Db 335 CGGGCTTGTACAGCTACTT 354

Search completed: March 13, 2005, 16:17:30
Job time : 98.2028 secs

THIS PAGE BLANK (USPTO)